

AD-A215 345

AGARD

ADVISORY GROUP FOR AEROSPACE RESEARCH & DEVELOPMENT

7 RUE ANCELLE 92200 NEUILLY SUR SEINE FRANCE

AGARD LECTURE SERIES No.165

**Modern Antenna Design Using
Computers and Measurement:
Application to Antenna Problems
of Military Interest**

S **D**
ELECTE
DEC 06 1989
B

NORTH ATLANTIC TREATY ORGANIZATION

DISTRIBUTION AND AVAILABILITY
ON BACK COVER

DISTRIBUTION STATEMENT A

Approved for public release;
Distribution Unlimited

89 12 05 258

AGARD-LS-165

NORTH ATLANTIC TREATY ORGANIZATION
ADVISORY GROUP FOR AEROSPACE RESEARCH AND DEVELOPMENT
(ORGANISATION DU TRAITE DE L'ATLANTIQUE NORD)

AGARD Lecture Series No. 165
**MODERN ANTENNA DESIGN USING
COMPUTERS AND MEASUREMENT:
APPLICATION TO ANTENNA PROBLEMS
OF MILITARY INTEREST**

This material in this publication was assembled to support a Lecture Series under the sponsorship of the Electromagnetic Wave Propagation Panel of AGARD and the Consultant and Exchange Programme of AGARD presented on 19–20 October 1989 in Ankara, Turkey, on 23–24 October 1989 in Lisbon, Portugal and on 26–27 October 1989 in London, United Kingdom.

THE MISSION OF AGARD

According to its Charter, the mission of AGARD is to bring together the leading personalities of the NATO nations in the fields of science and technology relating to aerospace for the following purposes:

- Recommending effective ways for the member nations to use their research and development capabilities for the common benefit of the NATO community;
- Providing scientific and technical advice and assistance to the Military Committee in the field of aerospace research and development (with particular regard to its military application);
- Continuously stimulating advances in the aerospace sciences relevant to strengthening the common defence posture;
- Improving the co-operation among member nations in aerospace research and development;
- Exchange of scientific and technical information;
- Providing assistance to member nations for the purpose of increasing their scientific and technical potential;
- Rendering scientific and technical assistance, as requested, to other NATO bodies and to member nations in connection with research and development problems in the aerospace field.

The highest authority within AGARD is the National Delegates Board consisting of officially appointed senior representatives from each member nation. The mission of AGARD is carried out through the Panels which are composed of experts appointed by the National Delegates, the Consultant and Exchange Programme and the Aerospace Applications Studies Programme. The results of AGARD work are reported to the member nations and the NATO Authorities through the AGARD series of publications of which this is one.

Participation in AGARD activities is by invitation only and is normally limited to citizens of the NATO nations.

The content of this publication has been reproduced directly from material supplied by AGARD or the authors.

Published September 1989

Copyright © AGARD 1989
All Rights Reserved

ISBN 92-835-0526-3



*Printed by Specialised Printing Services Limited
40 Chigwell Lane, Loughton, Essex IG10 3TZ*

THEME

The working environment in which an antenna is installed may substantially modify such antenna parameters as radiation efficiency, impedance, bandwidth, power handling capacity, pattern, etc. The need for more accurate antenna design, combined with the continuing growth of computational techniques, is complementing the more traditional approaches of measurement and analysis to vastly broaden the breadth and depth of problems that are now quantifiable. Computers are being used not only for numerical modelling/simulation, but also for measurement, data acquisition, and subsequent transformation of data. The newly available computational techniques are changing the way we think about, formulate, solve and interpret problems.

This Lecture Series, sponsored by the AGARD Electromagnetic Wave Propagation Panel, has been implemented by the Consultant and Exchange Programme.

* * *

L'environnement opérationnel peut modifier de façon significative certains paramètres d'une antenne tels que son rendement de rayonnement, son impédance, sa bande passante, sa puissance, son diagramme etc.

Le besoin d'une étude plus précise de la conception de l'antenne, combiné avec le développement continu des techniques de calcul viennent en complément aux approches plus traditionnelles de mesure et d'analyse pour élargir et approfondir les problèmes qui sont aujourd'hui quantifiables.

Les ordinateurs sont utilisés non seulement pour la modélisation et pour la simulation numériques, mais aussi pour la mesure, pour l'acquisition de données et pour le traitement ultérieur des données. Les techniques de calcul disponibles aujourd'hui sont en train de transformer notre manière de penser, de formuler, de résoudre et d'interpréter les problèmes.

Ce Cycle de Conférences est organisé dans le cadre du programme des Consultants et des échanges, sous l'égide du Panel AGARD sur la Propagation des Ondes Electromagnétiques.



Accession For	
NTIS GRA&I	<input checked="" type="checkbox"/>
DTIC TAB	<input type="checkbox"/>
Unannounced	<input type="checkbox"/>
Justification	
By _____	
Distribution/	
Availability Codes	
Dist	vail and/or Special
A-1	

LIST OF SPEAKERS

Lecture Series Director: Dr J.S.Belrose
POB 11490
Communications Research Center
Station H
Ottawa, Ontario K2H 8S2
Canada

AUTHORS/SPEAKERS

Dr Bruno Audone
Aeritalia
Gruppo Equipaggiamenti
Caselle Torinese 10072
Torino
Italy

Dr G.J.Burke
POB 5504
L-156 Lawrence Livermore National Lab.
Livermore, CA 94550
USA

Prof. Stanley Kubina
Electrician & Compatibility Engine
Concordia University
Loyola Campus
7141 Sherbrooke St W
Montreal, QC H4B 1R6
Canada

Prof. Dr Ing. Heinz Lindenmeier
Lehrstuhl für Hochfrequenztechnik
Hochschule der Bundeswehr München
Werner Heisenberg Weg 39
8014 Neubiberg
Germany

Dr Edmund K. Miller
General Research Corp.
5383 Hollister Ave.
Santa Barbara, CA 93111
USA

Mr Fred Molinet
Société Mothesim
La Boursidière
RN 186
92357 Le Plessis Robinson
France

Mr T.J. Murphy
Antenna Design Manager
BAe PLC — Dynamics Group
FPC 77
P O Box 5
Filton, Bristol BS12 7QW
United Kingdom

ABSTRACT

The working environment in which an antenna is installed may substantially modify such antenna parameters as radiation efficiency, impedance, bandwidth, power handling capacity, pattern, etc. The need for more accurate antenna design, combined with the continuing growth of computational techniques, are complementing the more traditional approaches of measurement and analysis to vastly broaden the breadth and depth of problems that are now quantifiable. Computers are being used not only for numerical modelling/simulation, but also for measurement, data acquisition, and subsequent transformation of data. The newly available computational techniques are changing the way we think about, formulate, solve and interpret problems.

The subject to be addressed is that of modern antenna design using computers and measurement, with emphasis on application to antenna problems of military interest. Computer models, using Moment Method and GTD computer codes, complemented by scale model measurement, and measurement in full scale can provide the essential performance information on which design decisions can be based, over the broad frequency band VLF to EHF. — N110 F111

The presentation will begin by a selective survey of computational electromagnetics for antenna application. This will be followed by a brief history of the successive development of the theory and recent advances, and by a discussion of a generalized signal-processing approach called model-based parameter estimation for improving the efficiency of electromagnetic modelling. The emphasis will be an application to problems of military interest. Environmental effects such as earth-air-interface effects on radiation and scattering (antennas on cliffs or antennas near the ground); land vehicular antennas, antennas on ships and aircraft; radar cross-section measurements and theory, which represent a key point in stealth technology; scale modelling, full scale measurement, and the compact ranges of antennas and RCS measurements are topics that will be addressed by a team of lecturers chosen to have expertise in the various areas relevant to the theme of the Lecture Series.

This Lecture Series sponsored by the AGARD Electromagnetic Wave Propagation has been implemented by the Consultant and Exchange Programme.

Appreciation is expressed to all who assisted in the organization of this Lecture Series, as well as the compilation of this volume, to the lecturers, AGARD staff and other collaborators and to those who will attend the LS, since lectures must be given to an audience and it is the interested response of the participants that makes the LS successful.

John S. Belrose
Lecture Series Director
April 1989

CONTENTS

	Page
THEME	iii
LIST OF SPEAKERS	iv
ABSTRACT	v
	Reference
INTRODUCTION by J.S.Belrose	1
A SELECTIVE SURVEY OF COMPUTATIONAL ELECTROMAGNETICS FOR ANTENNA APPLICATIONS by E.K.Miller	2
RECENT ADVANCES TO NEC: APPLICATIONS AND VALIDATION by G.J.Burke	3
SOME APPLICATIONS OF MODEL-BASED PARAMETER ESTIMATION IN COMPUTATION ELECTROMAGNETICS by E.K.Miller and G.J.Burke	4
MEASUREMENT AND COMPUTER SIMULATION OF ANTENNAS ON SHIPS AND AIRCRAFT FOR RESULTS OF OPERATIONAL RELIABILITY by S.J.Kubina	5
HF WIRE ANTENNAS OVER REAL GROUND: COMPUTER SIMULATION AND MEASUREMENT by J.S.Belrose, G.M.Royer and L.E.Petrie	6
ANTENNA ON LAND VEHICLES FOR IMPROVED COMMUNICATIONS by H.K.Lindenmeier	7
GTD/UTD: BRIEF HISTORY OF SUCCESSIVE DEVELOPMENT OF THEORY AND RECENT ADVANCES – APPLICATIONS TO ANTENNAS ON SHIPS AND AIRCRAFT by F.A.Molinet	8
COMPACT RANGES IN ANTENNA AND RCS MEASUREMENTS by B.Audone	9
ANTENNA INSTALLATION ON AIRCRAFT: THEORY AND PRACTICE by T.J.Murphy	10
BIBLIOGRAPHY	B

Introduction

John S. Belrose
Communications Research Centre
Ottawa ON K2H 8S2

The electromagnetic antenna is a most important element of a radio communications system since -no matter how sophisticated the communications system is, -how adaptive it is to varying propagation conditions, -how immune it is to interference, jamming and noise, -the ability to communicate successfully is often determined by the operational performance of the antennas used. It is the antenna that couples the signal to be transmitted into the propagation channel, and it is the antenna that probes the channel, receives the electromagnetic waves and feeds this information signal to the receiving equipment. Since the propagation channel is not lossless, or distortion free, it is important to start out and end as efficiently as possible, viz. to couple as much power as possible into the dominant propagation mode, and to receive as efficiently as possible the available power incident on the receiving antenna.

The electromagnetic antenna is perhaps the most misunderstood element of a radio communications system. Antenna engineers themselves will argue about how an antenna works or does not work, and when they conduct an experiment they are sometimes surprised about the result, or having conducted one experiment they find that this only leads to a requirement to conduct another and another, to clarify ones understanding. Even antenna companies, whose business it is to -develop antennas, -manufacture antennas, -promote the proper operational use of and market antennas, -often do not have a full understanding of, or in some cases provide misleading information and advice about the operational performance of the products they market. The working environment in which an antenna is installed may substantially modify such antenna parameters as: radiation efficiency, impedance, bandwidth, power handling capacity, radiation pattern, and etc. Recently developed computational techniques are complementing the more traditional approaches of measurement and analysis to vastly broaden the breadth and depth of antenna problems that can now be solved.

Computational electromagnetics is taking its place as an area of interest side-by-side with theoretical analysis and measurement. In addition micro- and personal-computers are increasingly used: -to tune antennas; -adapt their parameters in accord with varying propagation, noise and interference conditions or operational requirements; -and measure their characteristics; -beginning with the initial measurements (which nowadays employ computer controlled instrumentation) through comparison of measurements with computation, plotting the final result for viewing and interpretation, and preparation for camera ready figures. The purpose of this lecture series is to take a look at recent developments in the above fields. Of particular recent interest is the validation of computer codes by comparison with measurement or analysis (for the few cases where analysis can be done very accurately). A recent special issue of the Applied Computational Electromagnetics Society Journal [1989] has recently been published which addresses the subject of computer code validation.

In 1983, almost 6-years ago to the day, the Electromagnetic Wave Propagation Panel and the Consultant and Exchange Programme of AGARD presented a lecture series on the performance of antennas in their operational environments [AGARD LS,1983]. Some of the same speakers at that lecture series have prepared presentations for this one. In part the present lecture series can be viewed as a much needed update of the material/information presented in that earlier one; however, since considerable progress has been made in the field during the past 6-years, computational electromagnetics techniques have been expanded from investigations of the MOM technique to considerations also of the GTD and UTD techniques. The latter techniques provide the tools for modelling/simulation of antenna problems at higher frequencies.

The subject to be addressed is that of modern antenna design using computers and measurement: with emphasis on application to antenna problems of military interest. Computer codes complemented by scale model measurement, and measurement in full scale can provide the essential performance information on which design decisions can be based, over the broad frequency band VLF to EHF.

References

The performance of Antennas in their Operational Environments (edited by J.S. Belrose), AGARD LS No. 131, September 1983. Available: NTIS (Ref. NASA Acces. No. N84-12367).

Electromagnetics Computer Code Validation, Special Issue of Applied Computational Electromagnetics Society Journal, Naval Post Graduate School, Code 62AB, Monterey, CA 93943, 1989.

A Selective Survey of Computational Electromagnetics for Antenna Applications

E. K. Miller
General Research Corporation
5383 Hollister Avenue
Santa Barbara, CA 93111

ABSTRACT

The continuing growth of computing resources is changing how we think about, formulate, solve, and interpret problems. In electromagnetics as elsewhere, computational techniques are complementing the more traditional approaches of measurement and analysis to vastly broaden the breadth and depth of problems that are now quantifiable. An attempt is made in this article to place into perspective some of the tools used in computational electromagnetics with respect to the different kinds of approaches that may be used and their computer-resource requirements, paying particular attention to numerical models based on integral and differential equations. After a brief background discussion in Section 2, we review in Sections 3 and 4 respectively some of the analytical and numerical issues involved in developing a computer model. In Section 5 we include some practical considerations from the viewpoint of computer-resource requirements, followed by a discussion in Section 6 of some ways by which computer time might be reduced. Our presentation concludes in Section 7 with a brief examination of validation and error checking. Emphasis throughout is on review and summarization rather than detailed exposition.

I. INTRODUCTION

Computational Electromagnetics (CEM) may be broadly defined to be that branch of electromagnetics that intrinsically and routinely involves using a digital computer to obtain numerical results. With the evolutionary development of CEM during the past 20-plus years, two basic lines of improvement can be identified. One is due to advances taking place in computer hardware and software, providing tools of steadily growing power with little effort on the part of the electromagnetics community per se. The other line of improvement originates from within the electromagnetics discipline itself, where increasing awareness and utilization of numerical techniques has provided an expanding base of capability for solving problems in electromagnetics. The result has been to add the third tool of computational methods in EM specifically, and in science and engineering generally, to the two classical tools of experimental observation and mathematical analysis.

The goal of this article is to review concisely and conceptually some of the basic issues involved in CEM, to survey present capabilities and to contemplate future directions where appropriate. We attempt to accomplish this extremely broad task by introducing only the detail needed to illustrate the central ideas involved, and providing a selection of references from which the interested reader may obtain more information. Emphasis will be focussed therefore on the underlying principles which unify the various modeling approaches used in electromagnetics while avoiding most of the detail that makes them different. In Section 2 we begin with a brief discussion of the basic idea. In Sections 3 and 4 respectively, we examine some of the analytical and numerical issues involved in developing an electromagnetic computer model. Some of the practical implications of developing models in terms of the required computer resources are considered in Section 5, followed in Section 6 by a more detailed discussion of various means by which computer-time requirements might be reduced. In Section 7 we examine errors encountered and model validation, followed by a concluding summary in Section 8. While we attempt to provide some perspective concerning the relationship between differential- (DE) and integral-equation (IE) modeling, more attention will be devoted to the latter.

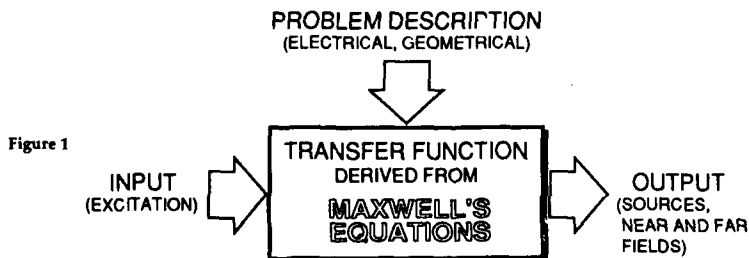
Listed throughout are representative, but not exhaustive, numbers of references which deal with various specialty aspects of CEM. For readers interested in broader, more general expositions, the well-known book on the Moment Method by Harrington (1968), the books edited by Mittra (1973, 1975), Uslenghi (1978) and Strait (1980), and the monographs by Stutzman and Thiele (1981), Popovic, Dragovic and Djordjevic (1982) and by Moore and Pizer (1984), are recommended.

II. BACKGROUND DISCUSSION

Electromagnetics is the scientific discipline that deals with electric and magnetic sources and the fields these sources produce in specified environments. Maxwell's Equations provide the starting point for the study of electromagnetic problems, together with certain principles and theorems such as superposition, reciprocity, equivalence, induction, duality, linearity, uniqueness, etc. derived therefrom [Harrington (1961), Van Nostrand Scientific Encyclopedia (1976)]. While a variety of specialized problems can be identified, a common ingredient of essentially all of them is that of establishing a quantitative relationship between a cause (forcing function or input) and its effect (the response or output), a relationship which we refer to as a field propagator, the computational characteristics of which are determined by the mathematical form used to describe it.

A. Modeling as a Transfer-Function

This relationship may be viewed as a generalized transfer function (see Fig. 1) in which two basic problem types become apparent. For the analysis or direct problem, the input is known and the transfer function is derivable from the problem specification, with the output or response to be determined. For the case of the synthesis or inverse problem, two problem classes may be identified. The "easier" synthesis problem involves finding the input given the output and transfer function, an example of which



is that of determining the source voltages which produce an observed pattern for a known antenna array. The "more difficult" synthesis problem itself separates into two problems. One is that of finding the transfer function, given the input and output, an example of which is that of finding a source distribution that produces a given far field. The other and still more difficult, is that of finding the object geometry which produces an observed scattered field from a known exciting field. The latter problem is the most difficult of the three synthesis problems to solve because it is intrinsically transcendental and nonlinear. Furthermore, such problems are subject to uniqueness constraints which can impose difficulties and uncertainties in developing their solutions.

Electromagnetic transfer functions are derived from a particular solution of Maxwell's Equations in the form of a field propagator, as the "cause" mentioned above normally involves some specified or known excitation whose "effect" is to induce some to-be-determined response. It therefore follows that the essence of electromagnetics is the study and determination of field propagators to thereby obtain an input-output transfer function for the problem of interest, and it follows that CEM inevitably does also. This observation, while perhaps appearing transparent, is actually an extremely fundamental one as it provides a focus for what CEM is all about, and provides a basis for classification of model types as we now discuss.

B. Some Issues Involved in Developing a Computer Model

In order to establish an appropriate perspective for subsequent discussion, we briefly consider here a classification of model types, the steps involved in developing a computer model, the desirable attributes of a computer model, and finally the role of approximation throughout the modeling process.

1. Classification of Model Types

It is convenient to classify solution techniques for electromagnetic modeling in terms of the field propagator that might be used, the anticipated application, and the problem type for which the model is intended to be used, as is outlined in Table 1. Selection of a field propagator in the form, for example, of the Maxwell curl equations, a Green's function, modal or spectral expansions, or an optical description, is a necessary first step in developing a solution to any electromagnetic problem.

2. Development of a Computer Model

Development of a computer model in electromagnetics or literally any other disciplinary activity can be decomposed into a small number of basic, generic steps. These steps might be described by different names, but would include at a minimum those outlined in Table 2. Note that by its nature, validation is an open ended process which cumulatively can absorb more effort than all the other steps together. The primary focus of the discussion which follows in this article is on the issue of numerical implementation.

3. Desirable Attributes of a Computer Model

A computer model must have some minimum set of basic properties to be useful. From the long list of attributes that might be desired, we consider: 1) accuracy, 2) efficiency, and 3) utility the three most important as summarized in Table 3. Accuracy is put foremost since results of insufficient or unknown accuracy have uncertain value and may even be harmful. On the other hand, a code that produces accurate results but at unacceptable cost will have hardly any more value. Finally, a code's applicability in terms of the depth and breadth of the problems for which it can be used determines its utility.

4. The Role of Approximation

As approximation is an intrinsic part of each step involved in developing a computer model, we summarize some of the more commonly used approximations in Table 4. We note that the distinction between an approximation at the conceptualization step and during the formulation is somewhat arbitrary, but choose to use the former category for those approximations that occur prior to the formulation itself.

III. ANALYTICAL ISSUES IN DEVELOPING A COMPUTER MODEL

As mentioned above, selection of a field propagator is a first step in developing an electromagnetic computer model. Further attention here is limited to propagators which employ either the Maxwell curl equations or source integrals which employ a Green's function. First we consider briefly selection of the solution domain and then selection of the field propagator with discussion here limited to integral- and differential-equation models.

A. Selection of Solution Domain

Either the IE or DE propagator can be formulated in the time domain, where time is treated as an independent variable, or in the frequency domain where the harmonic time variation $\exp(i\omega t)$ is assumed. The numerical treatment of generic time- and frequency-domain operators denoted by $L(\omega)$ and $L(t)$ respectively is compared in Table 5.

Whatever propagator and domain are chosen, the analytically formal solution can be numerically quantified via use of the method of moments (MoM) [Harrington (1968)], leading ultimately to a linear system of equations as a result of developing a discretized and sampled approximation to the continuous (generally) physical reality being modeled. Developing the approach that may be best suited to a particular problem involves making tradeoffs among a variety of choices throughout the analytical formulation and numerical implementation. In the following discussion, we consider some aspects of these choices and their influence on the utility of the computer model which eventually results.

B. Selection of Field Propagator

We briefly discuss and compare below the characteristics of the various propagator-based models in terms of their development and applicability.

1. Integral-equation (IE) model

The basic starting point for developing an IE model in electromagnetics is selection of a Green's function appropriate for the problem class of interest. While there are a variety of Green's functions from which to choose, a typical starting point for most IE MoM models is that for an infinite medium. Although the formulation might be accomplished in various ways, one of the more straightforward is based on the scalar Green's function and Green's theorem. This leads to the Kirchhoff integrals

[Stratton (1941), p. 464, et. seq.] from which the fields in a given contiguous volume of space can be written in terms of integrals over the surfaces which bound it and volume integrals over those sources located within it. Exceptions to this general rule are discussed below in Section 4.

Analytical manipulation of a source integral which incorporates the selected Green's function as part of its kernel function then follows, with the specific details depending on the particular formulation being employed. Perhaps the simplest is that of boundary-condition matching wherein the behavior required of the electric and/or magnetic fields at specified surfaces which define the problem geometry is explicitly imposed. Alternative formulations, for example, the Rayleigh-Ritz variational method and Rumsey's reaction concept might be used instead, but as pointed out by Harrington (1980), from the viewpoint of a numerical implementation any of these approaches lead to formally equivalent models.

This analytical formulation leads to an integral operator, whose kernel can include differential operators as well, which acts on the unknown source or field. Although it would be more accurate to refer to this as an integro-differential equation, it is usually called simply an integral equation. Two kinds of integral equations are obtained, one known as a Fredholm integral equation of the first kind in which the unknown appears only under the integral, and the other a second-kind equation in which the unknown also appears outside the integral. In Table 6 are included some of the integral equations used in CEM, including the magnetic-field integral equation (MFIE) and electric-field integral equation (EFIE) for perfect conductors for both the frequency domain and time domain.

2. Differential-Equation (DE) Model

A DE MoM model, being based on the defining Maxwell's Equations, requires intrinsically less analytical manipulation than does derivation of an IE model. Numerical implementation of a DE model, however, can differ significantly from that used for an IE formulation in a number of ways for several reasons:

- i. The differential operator is a local rather than global one in contrast to the Green's function upon which the integral operator is based. This means that the spatial variation of the fields must be developed from sampling in as many dimensions as possessed by the problem, rather than one less as the IE model permits if an appropriate Green's function is available.
- ii. The integral operator includes an explicit radiation condition.
- iii. The differential operator includes a capability to treat medium inhomogeneities, non-linearities, and time variations in a more straightforward manner than does the integral operator.

These and other differences between development of IE and DE models are summarized in Table 7, with their modeling applicability compared in Table 8.

3. Modal-Expansion Model

Modal expansions are useful for propagating electromagnetic fields because the source-field relationship can be expressed in terms of well-known analytical functions as an alternate way of writing a Green's function for special distributions of point sources. In two dimensions for example, the propagator can be written in terms of circular harmonics and cylindrical Hankel functions. Corresponding expressions in three dimensions might involve spherical harmonics, spherical Hankel functions and Legendre polynomials. Expansion in terms of analytical solutions to the wave equation in other coordinate systems can also be used, but requires computation of special functions that are generally less easily evaluated, such as Mathieu functions for the 2D solution in elliptical coordinates and spheroidal functions for the 3D solution in oblate or prolate spheroidal coordinates.

One implementation of modal propagators for numerical modeling is that due to Waterman (1965, 1973) whose approach employs the extended boundary condition (EBC) whereby the required field behavior is satisfied away from the boundary surface on which the sources are located. This procedure, widely known as the "T-Matrix" approach has evidently been more widely employed in optics and acoustics than in electromagnetics [Barber, et. al. (1982), Varadan and Varadan (1982)]. In what amounts to a reciprocal application of EBC, the sources are removed from the boundary surface on which the field-boundary conditions are applied as reported by Komiya (1984), Ludwig (1986), Leviatan and Boag (1987) and others. These modal techniques seem to offer some computational advantages for certain kinds of problems and might be regarded as employing entire-domain basis and testing functions, but nevertheless lead to linear systems of equations whose numerical solution is required. Fourier-transform solution techniques might also be included in this category [Kastner and Mittra (1983a), (1983b)], since they do involve modal expansions, but that is a specialized area which we do not pursue further here.

4. Geometrical-Optics Model

Geometrical optics and the geometrical theory of diffraction (GTD) are high-frequency asymptotic techniques wherein the fields are propagated using optical concepts such as shadowing, ray tubes, refraction and diffraction, etc. Although conceptually straightforward, optical techniques are limited by the unavailability of diffraction coefficients for various geometries and material bodies, and by the need to trace rays over complex surfaces. There is a vast literature on geometrical optics and GTD as may be ascertained by examining the yearly and cumulative indexes of such publications as the *IEEE Antennas and Propagation Society*. We devote no further attention to optics techniques here except for their combination with other propagators to develop hybrid models (see section VIA.2 below).

IV. NUMERICAL ISSUES IN DEVELOPING A COMPUTER MODEL

A. Sampling functions

At the core of numerical analysis is the idea of polynomial approximation, an observation made by Arden and Astill (1969) in facetiously using the subtitle "Numerical Analysis or 1001 Applications of Taylor's Series". The basic idea is to approximate quantities of interest in terms of sampling functions, often polynomials, that are then substituted for these quantities in various analytical operations. Thus, integral operators are replaced by finite sums and differential operators are similarly replaced by finite differences. For example, use of a first-order difference to approximate a derivative of the function $F(x)$ in terms of samples $F(x_1)$ and $F(x_2)$ leads to

$$dF(x)/dx \approx [F(x_2) - F(x_1)]/h; \quad x_1 \leq x \leq x_2 \quad (1a)$$

and implies a linear variation for $F(x)$ between x_1 and x_2 as does use of the trapezoidal rule

$$\int_{x_1}^{x_2} F(x) dx \approx h[F(x_1) + F(x_2)]/2 \quad (1b)$$

to approximate the integral of $F(x)$, where $h = x_i - x_{i-1}$. The central-difference approximation for the 2nd derivative, $d^2F(x)/dx^2 = [F(x_{i-1}) - 2F(x_i) + F(x_{i+1}))]/h^2$; (1c)
 similarly implies a quadratic variation for $F(x)$ around $x_0 = x_i - h/2 = x_i + h/2$, as does use of Simpson's rule

$$\int_{x_0}^{x_2} F(x) dx = h[F(x_0) + 4F(x_1) + F(x_2)]/6 \quad (1d)$$

to approximate the integral. Other kinds of polynomials and function sampling can be employed, as discussed in a large volume of literature, some examples of which are Abramowitz and Stegun (1964), Acton (1970), and Press et. al. (1986). It is interesting to see that numerical differentiation and integration can be accomplished using the same set of function samples and spacings, differing only in the signs and values of some of the associated weights. Note also that the added degrees of freedom which arise when the function samples can be unevenly spaced, as in Gaussian quadrature, produce a generally more accurate result (for well-behaved functions) for a given number of samples. This suggests the benefits that might be derived from using unequal sample sizes in MoM modeling should a systematic way of determining the best nonuniform sampling scheme be developed.

B. The Method of Moments (MoM)

Numerical implementation of the moment method is a relatively straightforward, and an intuitively logical extension, of these basic elements of numerical analysis, as described in the well-known book by Harrington (1968) and discussed and used extensively in CEM [see for example Mittra (1973, 1975), Strait (1980), Strait and Adams (1980), Harrington et. al. (1981), Perini and Buchanan (1985), Ney (1985), Itoh (1986), Poggio and Miller (1973), (1988)]. Whether it is an integral equation, a differential equation, or another approach that is being used for the numerical model, there are three essential sampling operations that are involved in reducing the analytical formulation via the moment method to a computer algorithm as outlined in Table 9. We note that operator sampling can ultimately determine the sampling density needed to achieve a desired accuracy in the source-field relationships involving integral operators, especially at and near the "self term" where the observation and source points become coincident or nearly so. Whatever the method used for these sampling operations, they lead to a linear system of equations or matrix approximation of the original integral or differential operators. Since the operations and choices involved in developing this matrix description is common to all moment-method models, we discuss them in somewhat more detail below.

When using IE techniques, the coefficient matrix in the linear system of equations which results is most often referred to as an impedance matrix because in the case of the E-field form, its multiplication of the vector of unknown currents equals a vector of electric fields or voltages. The inverse matrix similarly is often called an admittance matrix because its multiplication of the electric-field or voltage vector yields the unknown-current vector. In this discussion we instead use the terms direct matrix and solution matrix since they are more generic descriptions whatever the forms of the originating integral or differential equations. As illustrated below, development of the direct matrix and solution matrix dominates both the computer time and storage requirements of numerical modeling.

In the particular case of an IE model, the coefficients of the direct or original matrix are the mutual impedances of the multi-port representation which approximates the problem being modeled, and the coefficients of its solution matrix (or equivalent thereof) are the mutual admittances. Depending on whether a sub-domain or entire-domain basis has been used (see following section), these impedances and admittances represent either spatial or modal interactions among the N ports of the numerical model. In either case, these coefficients possess a physical relatability to the problem being modeled, and ultimately provide all the information available concerning any electromagnetic observables that are subsequently obtained.

Similar observations might also be made regarding the coefficients of the DE models, but whose multi-port representations describe local rather than global interactions. Because the DE model almost always leads to a larger, albeit less dense, direct matrix, its inverse (or equivalent) is rarely computed. It is worth noting that there are two widely used approaches for DE modeling, finite-difference (FD) and finite-element (FE) methods [Jeng and Chang (1984), Mason and Anderson (1985)]. They differ primarily in how the differential operators are approximated and the differential equations are satisfied, although the FE method commonly starts from a variational viewpoint while the FD approach begins from the defining differential equations. The FE method is generally better suited for modeling problems with complicated boundaries to which it provides a piecewise linear approximation as opposed to the cruder stairstep approximation of FD.

1. Factors involved in choosing basis and weight functions

Basis- and weight-function selection plays a critical role in determining the accuracy and efficiency of the resulting computer model. One goal of the basis- and weight-function selection is to minimize computer time while maximizing accuracy for the problem set to which the model is to be applied. Another, possibly conflicting goal, might be that of maximizing the collection of problem sets to which the model is applicable. A third might be to replicate the problem's physical behavior with as few samples as possible. Some of the generic combinations of bases and weights that are used for MoM models are listed below in Table 10 [Poggio and Miller (1973)].

a. Basis-function Selection

We note that there are two classes of bases used in MoM modeling, sub-domain and entire-domain functions. The former involves the use of bases which are applied in a repetitive fashion over sub-domains or sections (segments for wires, patches for surfaces, cells for volumes) of the object being modeled. The simplest example of a sub-domain basis is the single-term basis given by the "pulse" or stairstep function, which leads to a single, unknown constant for each sub-domain. Multi-term bases involving two or more functions on each sub domain and an equivalent number of unknowns are more often used for sub-domain expansions.

The entire-domain basis, on the other hand, uses multi-term expansions extending over the entire object, for example a circular harmonic expansion in azimuth for a body of revolution. As for sub-domain expansions, an unknown is associated with each term in the expansion. Examples of hybrid bases can also be found, where sub-domain and entire-domain bases are used on different parts of an object [Bornholdt and Medgyesi-Mitschang (1986)].

Although sub-domain bases are probably more flexible in terms of their applicability, they have a disadvantage generally not exhibited by the entire-domain form, which is the discontinuity that occurs at the domain boundaries. This discontinuity arises because an n_i -term sub-domain function can provide at most n_i-1 th continuity to an adjacent basis of the unknown it represents

assuming one of the n_i constants is reserved for the unknown itself. For example, the three-term or sinusoidal sub-domain basis $a_i + b_i \sin(ks) + c_i \cos(ks)$ used for wire modeling can represent a current continuous at most up to its first derivative. This provides continuous charge density, but produces a discontinuous first derivative in charge equivalent to a tri-pole charge at each junction.

As additional terms are used to develop a sub-domain basis, higher-order continuity can be achieved in the unknown that the basis represents assuming still that one constant is reserved for the unknown. In the general case of the n_i -term sub-domain basis, up to $n_i - 1$ constants can be determined from continuity conditions with the remainder reserved for the unknown. The kind of basis function employed ultimately determines the degree of fit that the numerical result can provide to the true behavior of the unknown for a given order of matrix. An important factor that should influence basis-function selection then is how closely a candidate function might resemble the physical behavior of the unknown it represents. Another consideration is whether a system of equations that is numerically easier to solve might result from a particular choice of basis and weight function, for example by increasing its diagonal dominance so that an iterative technique will converge more rapidly.

b. Weight-function Selection

The simplest weight that might be used is a delta function which leads to a point-sampled system of equations. But point sampling of the field operators can reveal any numerical anomalies that might arise as a result of basis-function discontinuities. Distributed, multi-term weight functions can also be employed on either a sub-domain or an entire-domain basis to provide a further smoothing of the final equations to be solved. One example of this is the special case where the same functions are used for both the bases and weights, a procedure known as Galerkin's method. The kind of testing function employed ultimately determines the degree to which the equations can be matched for a given basis function and number of unknowns. Some specific examples of basis- and weight-function combinations used in electromagnetics are summarized in Table 11.

2. Computing the Direct Matrix

We observe that obtaining the coefficients of the direct matrix in IE modeling is generally a two-step process. The first step is that of evaluating the defining integral operator in which the unknown is replaced by the basis functions selected. The second step involves integration of this result multiplied by the weight function selected. When using delta-function weights this second step is numerically trivial. But when using non-delta weights, such as the case in a Galerkin approach where the same function is used for both basis and weights, this second step can be analytically and numerically challenging.

Among the factors affecting the choice of the basis and weight functions therefore, one of the most important is that of reducing the computational effort needed to obtain the coefficients of the direct matrix. This is one of the reasons, aside from their physical appeal, why sinusoidal bases are often used for wire problems. In this case, where piecewise linear, filamentary current sources are most often used in connection with the thin-wire approximation, field expressions are available in easily evaluated, analytical expressions [Richmond (1965), Miller and Deadrick (1975)]. This is the case as well where Galerkin's method is employed [Richmond (1974)].

Aside from such special cases however, numerical evaluation of the direct matrix coefficients will involve the equivalent of point sampling of whatever order is needed to achieve the desired accuracy as illustrated below. Using a wirelike one dimensional problem to illustrate this point, we observe that at its most elementary level evaluation of the i 'th matrix coefficient then involves evaluating integrals of the form

$$\begin{aligned} Z_{ii} &= \int w_i(s) \int [b(s') K(s, s')] ds' ds \\ &= \sum p_m q_n w_i(s_m) b_n(s'_m) K(s_m, s'_m) \\ &= \sum \sum p_m q_n z(i, j, m, n); m=1, \dots, M(i, j) \end{aligned} \quad \begin{matrix} n=1, \dots, N(i, j) \\ i, j=1, \dots, N \end{matrix} \quad (2)$$

where $K(s, s')$ is the IE kernel function, and s_m and s'_m are the n 'th and m 'th locations of the observation and source integration samples. Thus, the final, direct-matrix coefficients can be seen to be "constructed" from sums of the more elementary coefficients $z(i, j, m, n)$ weighted by the quadrature coefficients p_m and q_n used in the numerical integration, which will be the case whenever analytical expressions are not available for the Z_{ij} . These elementary coefficients, given by $w_i(s_m) b_n(s'_m) K(s_m, s'_m)$, can in turn be seen to be simply products of samples of the integral-equation kernel or operator, and sampled basis and testing functions. It should be apparent from this expanded expression for the direct-matrix coefficients that interchanging the basis and weight functions leaves the final problem description unchanged, although the added observation that two different integral equations can yield identical matrices when using equivalent numerical treatments is less obvious [Wilton and Butler (1976), (1981)].

3. Computing the Solution Matrix

Once the direct matrix has been computed, the solution can be obtained numerically using various approaches. These range from inversion of the direct matrix to developing a solution via iteration as summarized in Table 12. A precautionary comment is in order with respect to the accuracy with which the solution matrix might be obtained. As computer speed and storage have increased, the number of unknowns employed in modeling has also increased, from a few 10s in earlier years to 1,000s now when using IE techniques. The increasing number of operations involved in solving these larger matrices increases sensitivity of the results to roundoff errors. This is especially the case when the direct matrix is not well-conditioned. It is therefore advisable to perform some sensitivity analyses to determine the direct-matrix condition number and to ascertain the possible need for performing some of the computations in double precision.

4. Obtaining the Solution

When a solution matrix has been developed using inversion or factorization, subsequently obtaining the solution is computationally straightforward, involving multiplication of the right-hand-side (RHS) source vector by the solution matrix. When an iterative approach is used, a solution matrix is not computed but the solution is instead developed from RHS-dependent manipulation of the direct matrix. Motivation for the latter comes from the possibility of reducing the N_i^3 dependency of the direct procedure. As problem size increases, the computation cost will be increasingly dominated by the solution time.

V. SOME PRACTICAL CONSIDERATIONS

Although the overall solution effort has various cost components, perhaps the one most considered is the computer time and storage required to obtain the numerical results desired. With the increasing computer memories becoming available, where even

micro-computers and workstations can directly address giga-bytes, the memory costs of modeling are becoming generally less important than the time cost, with which we are primarily concerned here. For each model class considered, the computer-time dependence on the number of unknowns is presented in a generic formula followed by the highest-order ($L/\Delta L$) term in that formula to demonstrate how computer time grows with increasing problem size (see Tables VII and XIII for more details).

A. Integral-equation modeling

1. Frequency domain

If we consider an IE model specifically, we can show that in general, the computer time associated with its application is dependent on the number of unknowns N_u in the frequency domain as

$$T_{IE,w} = A_{dir} N_u^2 + A_{solve} N_u^3 + A_{source} N_u^2 N_{rhs} + A_{field} N_u N_{rhs} N_{fields} \propto (L/\Delta L)^{2(D-1)} \quad (3a)$$

where the A's are computer- and algorithm-dependent coefficients which account for computation of:

A_{dir} —the direct (impedance) matrix,

A_{solve} —the solution (admittance) matrix (assuming inversion or factorization),

A_{source} —the source response (currents and charges) for one of N_{rhs} different excitations or right-hand sides (the g term of Table V),

A_{field} —one of N_{fields} fields, where $A_{field} \leq A_{dir}$ depending on whether a near-field (=) or far-field (<) value is obtained,

D is the problem dimensionality (for a wire IE model, $D = 2$ except when used for wire-mesh approximations of surfaces in which case $D = 3$), L is a characteristic length of the object being modeled and ΔL is the spatial resolution required, being proportional to the wavelength.

2. Time domain

A similar relationship holds for a time-domain IE model which uses N_t time steps,

$$T_{IE,t} = A_{dir} N_u^2 + A_{solve} N_u^3 + A_{source} N_u^2 N_{rhs} + A_{field} N_u N_{rhs} N_{fields} \propto (L/\Delta L)^{2(D-1)+r}, \quad \text{explicit approach, } 0 \leq r \leq 1 \\ \propto (L/\Delta L)^{2(D-1)}, \quad \text{implicit approach} \quad (3b)$$

with the A's accounting for computation of the time-domain terms equivalent to their frequency-domain counterparts above (with different numerical values), and r including the right-hand-side dependency. Although a direct matrix may require solution initially prior to time-stepping the model, that is normally avoided by using $\delta t \leq \Delta x/c$, which yields an explicit solution in which case $A_{solve} = 0$. As can be appreciated from these expressions, the number of unknowns that is required for these computations to be acceptably accurate has a strong influence on the computer time eventually needed.

B. Differential-equation modeling

1. Frequency domain

DE modeling is less commonly used in the frequency domain primarily because the order of the matrix that results depends on $(L/\Delta L)^D$ rather than the usual $(L/\Delta L)^{D-1}$ dependency of an IE model. On the other hand, the matrix coefficients require less computation whether the DE model is based on a finite-difference or finite-element treatment. Furthermore, the matrix is very sparse because a differential operator is a local rather than a global one as is the integral operator. Matrix fill time is therefore generally not of concern, and the overall computer time is given approximately by

$$T_{DE,w} = A_{solve} N_u W^2 + A_{source} N_u W N_{rhs} + A_{field} N_u^{(D-1)/D} N_{rhs} N_{fields} \propto (L/\Delta L)^{D-2} \quad (3c)$$

which exhibits a dominance by the matrix-solution term. Note that the banded nature of the DE direct matrix has been taken into account where the bandwidth W varies as N_u^0 , $N_u^{1/2}$, and $N_u^{1/3}$ respectively ($N_u^{(D-1)/D}$) in 1, 2, and 3 dimensions.

2. Time domain

Time-domain DE modeling can employ either implicit or explicit solution methods for developing the time variation of the solution. An explicit technique is one whereby the update at each time step is given in terms of solved-for past values of the unknowns and the present excitation, with no interaction permitted between unknowns within the same time step. An implicit technique on the other hand does allow for interaction of unknowns within the same time step, but can therefore require the solution of a matrix equation. In spite of this disadvantage, implicit techniques are important because they are not subject to Courant instability when $\delta t > \Delta x$ as is an explicit approach.

The solution time for the explicit case is approximated by

$$T_{DE,t} = A_{source} N_u N_{rhs} + A_{field} N_u^{(D-1)/D} N_{rhs} N_{fields} \propto (L/\Delta L)^{D-1+r}, \quad \text{explicit approach; } 0 \leq r \leq 1 \\ = A_{solve} N_u W^2 + A_{source} N_u W N_{rhs} + A_{field} N_u^{(D-1)/D} N_{rhs} N_{fields} \propto (L/\Delta L)^{D-2}, \quad \text{for } D = 2, 3 \\ \text{and } \propto (L/\Delta L)^{2+r}, \quad \text{for } D = 1, \text{ implicit approach; } 0 \leq r \leq 1 \quad (3d)$$

assuming a banded matrix is employed to solve the implicit direct matrix.

C. Discussion

It should be recognized that the above computer-time estimates assume solutions are obtained via matrix factorization, an N^3 process, and that iterative techniques when applicable should be expected to reduce the maximum order of the $(L/\Delta L)$ dependency, but at the cost however of requiring the computation to be repeated for each right-hand side. We also emphasize that these comparisons consider only problems involving homogeneous objects thereby providing a more favorable situation for IE models because their sampling dimensionality $d = D - 1$ for a problem dimensionality of D , but which increases to $d = D$ when an inhomogeneous object is modeled. Because of these and other factors which can lead to many different combinations of formulation and numerical treatment, the results above should be viewed as only generic guidelines, with the computational characteristics of each specific model requiring individual analysis to obtain numerical values for the various A_u coefficients and their $(L/\Delta L)$ dependency. It is relevant to observe that the lowest-order size dependency for three-dimensional problems is exhibited by the DE explicit time-domain model which is of order $(L/\Delta L)^4$.

An additional factor that should be considered when choosing among alternate computer models is the information needed for a particular application relative to the information provided by the model. A time-domain model for example can intrinsically provide a frequency response over a band of frequencies from a single calculation, whereas a frequency-domain model requires repeated evaluation at each of the frequencies required to define the wideband response (see Table V for discussion). Iterative solution of the direct matrix may be preferable for problems involving only one, or a few, excitations such as is the case for antenna modeling, to avoid computing all N^2 admittances of the solution matrix when only a single column of that matrix is needed. A DE-based model necessarily provides the "near" fields throughout the region being modeled, while an IE-based model requires additional computations essentially the same as those done in filling the impedance matrix once the sources have been obtained to evaluate the near fields. For applications which require modest computer time and storage, these considerations may be relatively less important than those which strain available computer resources. Clearly, the overall objective from an applications viewpoint is to obtain the needed information at the required level of accuracy for the minimum overall cost.

D. Sampling requirements

We may estimate the number of samples needed to adequately model the spatial, temporal, and angular variation of the various quantities of interest in terms of an object characteristic length L , and sampling dimension d . This may be done from knowledge of the typical spatial and temporal densities determined from computer experiments, and/or from invocation of Nyquist-like sampling rates for field variations in angle as a function of aperture size. The resulting estimates are summarized in Table 13 and apply to both IE and DE models.

These may be regarded as wavelength-driven sampling rates, in contrast with the geometry-driven sampling rates that can arise due to problem variations that are small in scale compared with L . Geometry-driven sampling would affect primarily N_s , resulting in larger values than those indicated above.

We note that the computer time is eventually dominated by computation of the solution matrix, and can grow as $(L/\Delta L)^2$, $(L/\Delta L)^3$ and $(L/\Delta L)^4$ (or P , P^2 , and P^3) respectively for wire, surface and volume objects modeled using integral equations and matrix factorization or inversion. Thus, in spite of the fact that mainframe computer power has grown by a factor of about 10^6 from the UNIVAC-1 to the CRAY2, the growth in problem size is much less as illustrated by Fig. 2. The curves on this graph demonstrate emphatically the need for finding faster ways of performing the computations, some aspects of which we next discuss.

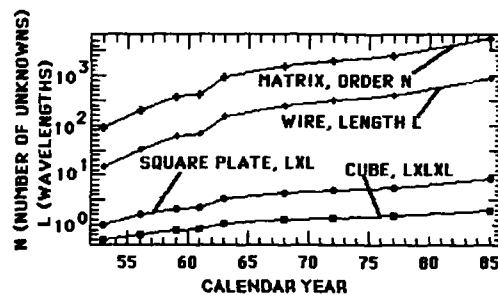


Figure 2. Time development of IE-based modeling capability for 1D (e.g., a wire), 2D (e.g., a plate), and 3D (e.g., a penetrable, inhomogeneous cube) sampling of a problem of characteristic dimension L in wavelengths, and matrix order N solvable in one hour of computer time using mainframe computers introduced in the years indicated. Linear-systems solution using LU decomposition (an N^3 dependency) is assumed with number of unknowns proportional to L , L^2 and L^3 respectively without any problem symmetry being exploited. These results should be viewed as upper bounds on solution time and might be substantially reduced by advances in linear-system solution procedures.

VI. WAYS OF DECREASING COMPUTER TIME

The obvious drawback of direct moment-method models as N_s increases with increasing problem size and/or complexity suggests the need for less computationally intensive alternatives. There are various alternatives for decreasing the computer cost associated with solving electromagnetic problems using the method of moments. The basic intent in any case is either to reduce the direct cost of performing a given modeling computation, or to reduce the number of modeling computations needed to obtain a desired result. These might include analytical, computational, and experimental approaches or combinations thereof, some of which we briefly consider here.

A. Analytical

Analytical alternatives might include extensions of the basic approach such as use of specialized Green's functions, combination of two or more approaches in a hybrid formulation, or a variety of other possibilities, some of which are briefly discussed below and outlined in Table 14.

1. Specialized Green's functions

The basic integral equations presented above employ an infinite-medium Green's function (IMGF), use of which leads to a problem description in terms of only the surface sources. The IMGF "propagates" the fields of these sources throughout the problem volume of interest, and satisfies as its only "boundary condition" the Sommerfeld radiation condition as $r \rightarrow \infty$. The required boundary conditions on the object(s) being modeled are satisfied through the MoM solution of the formal integral equation.

One analytical approach for developing more efficient IE models is use of the specialized Green's functions which satisfy additional boundary conditions. Although such Green's functions are available for only a relatively small set of separable geometries,

they nonetheless provide a useful alternative for some problems of practical interest. These include planes, cylinders, spheres, ellipsoids, etc. for which delta-source Green's functions provide formally a means of modeling wires or other objects located in their vicinity. The tradeoff in using such Green's functions is that of reducing the number of unknowns needed in the model at the expense of dealing with a significantly more complex integral-equation kernel.

For example, modeling an object near the infinite, planar interface between two half spaces could include the interface tangential fields as unknowns as well as those associated with the object itself [Miller (1981)]. Alternatively, a Green's function which satisfies the interface boundary conditions can be used instead, one possibility being inclusion of the Sommerfeld integrals as part of the integral-equation kernel [Miller, et. al. (1972a), (1972b)]. Although this is advantageous in that the fields over a (necessarily) truncated area of the infinite interface do not need explicit solution, the complexity of numerically evaluating the Sommerfeld integrals must then be dealt with. Even so, the advantages of using the specialized Green's function make it an attractive alternative. Examples are given by: Lucke (1951) and Lytle (1971) for a cylinder; by Einziger and Felsen (1983) for a cylindrical sector; by Tesche (1972) for a planar region bounded by perfectly conducting, parallel plates; by Tesche and Neureuther (1970) for a sphere; and by Pathak (1983) for an eigen-function expansion of dyadic Green's functions. In Table 15, we summarize several of the types of Green's functions that have been or might be used in IE modeling.

2. Hybrid approaches

One example of a hybrid approach is that of combining MoM with the geometrical theory of diffraction (GTD), as first reported by Thiele and Newhouse (1969) and more recently as summarized by Burnside and Pathak (1980). GTD has the attractive property that the computation time is essentially independent of increasing problem size or frequency, in contrast with the moment method. On the other hand, increasing complexity can result in the need to trace more rays and compute more diffraction coefficients, which can ultimately limit the applicability of GTD models. As an alternative to modeling a given problem using only the MoM or GTD alone, they can be used in combination to exploit their complementary advantages, thereby obtaining a model which is significantly better than either one used separately. Such "hybrid" techniques appear to be one of the more promising means for developing the models needed for large, complex problems.

Other hybrid techniques have been developed, one example being the combined use of a DE model and a modal expansion, termed the unimoment method by Mei (1974). As implemented by Mei, this technique involves solving a bounded, three-dimensional, body of revolution using a DE model and matching this solution across a spherical enclosing surface to an external, infinite medium where the solution is given by a modal expansion. A similar solution to a two-dimensional problem was presented by Miller and Olte (1966) for an inhomogeneously sheathed, infinite cylinder. Combination of DE and IE modeling in the time domain is described by Taflov and Umashankar (1980), where the IE model provides the short-circuit currents on a body whose interior response is subsequently obtained using the DE model. Several of these hybrid treatments are summarized in Table 16.

3. Other

Other analytically based methods for simplifying the formulation and/or the subsequent computation include the impedance boundary condition (IBC) and the physical optics (PO) approximation. The IBC [Rytov (1940), Leontovich (1948), Senior (1960, 1981), Mitzner (1967)] involves an assumption that the tangential electric and magnetic fields on the surface of a penetrable object can be related through a surface impedance Z_{surf} e.g., $E_{tan} = -Z_{surf} n \times H_{tan}$ where n is a surface normal and $Z_{surf} = (1 + j)\omega\mu\delta/2$ with δ the skin depth. This assumption halves the number of unknowns needed for modeling the object since the magnetic field can replace the electric field throughout the IE model, as is usually done, or vice versa. The benefit of the IBC occurs primarily in the computational effort needed in obtaining the solution matrix, as the direct-matrix computation is essentially unchanged because the interaction fields of the replaced unknown must still be included.

For acceptable results, the radius of curvature of the body and the body size in the direction of dominant current flow must be large with respect to skin depth δ . A modified form which is applicable to bodies having a smaller radius of curvature is given by $E_{tan} = (1 - p)Z_{surf}H_{tan}$ and $E_n = (1 + p)Z_{surf}H_n$ where $p = (1 - j)\delta(C_1 - C_2)/4$ and C_1 and C_2 are the principal curvatures [Mitzner (1967)]. Note that the object need not be solid, as the IBC can also be used to develop a sheet-impedance model for dielectric shells [Harrington and Mautz (1975)].

Just as the IBC can be used to eliminate one half of the unknown surface fields involved in modeling a penetrable body, the PO approximation can eliminate the remaining unknowns. It involves making the further assumption that the surface current can be expressed directly in terms of the incident field. For a perfect conductor, the substitution $J_s = 2n \times H^{inc}$ is made, although more generally we might use $J_s = (1 - R)n \times H^{inc}$ and $K_s = -(1 + R)n \times E^{inc}$, where R is the plane-wave reflection coefficient for the object. Using the PO approximation thus circumvents the need to compute a direct matrix or to obtain its solution, and the fields resulting from the incident excitation can be computed directly by evaluating the appropriate source integrals.

A further logical extension of IBC and PO is typified by the reflection-coefficient approximation (RCA) or modified-image theory that has been used in modeling interface problems [Miller, et. al. (1972a), (1972b), Burke and Miller (1984)]. In applying the RCA, not only is solution of the interface sources avoided, but their integration over the interface is circumvented by approximating their effects with image fields multiplied by Fresnel plane-wave reflection coefficients. This is an exact procedure for perfectly conducting, planar interfaces, but is otherwise an approximation although the RCA has a reasonably wide scope of applicability. The RCA is one way of avoiding the rigor, but computational complexity, of the Sommerfeld Green's function for the interface problem as discussed above. By extending the concept of an image treatment, an approach of rigor comparable to the Sommerfeld treatment can be developed [Lindell, et. al. (1985)], but also of comparable computational complexity [Burke and Miller (1986)].

B. Numerical

There are a number of numerical possibilities by which the cost of a moment-method model might be reduced. These include using alternate solution techniques for solving the linear system of equations and fully exploiting problem symmetries, the goal of either being to reduce the N^2 and N^3 dependencies of the matrix fill and solution terms in Eq. (3). Another possibility is provided by the so-called near-neighbor approximation (NNA) which involves neglecting interactions between source-field points separated by more than some specified distance when using a sub-domain basis [Ferguson, Lehman, and Balestri (1976)], or between modes separated by more than some wavenumber difference when using an entire-domain basis [Medgyesi-Mitschang and Putnam (1983), (1985)]. The fill time for the direct matrix and its subsequent solution time can then also be reduced, although

achieving the latter depends on being able to exploit the sparseness of a matrix whose nonzero coefficients may in general be widely distributed throughout it. Some of these numerical means of reducing computer time are summarized in Table 17.

1. Alternate Solution Techniques

The number of multiplies/divides (operations) needed in applying direct solution techniques to linear systems is proportional to N_s^3 . For this computational investment there is obtained a solution matrix which is independent of the RHS forcing function or excitation. Solutions are then available for an additional computational cost per solution proportional to N_s^2 (multiplication of an N_s 'th order matrix by an N_s 'th order vector or the equivalent).

If a RHS-dependent solution is acceptable, then techniques which have a lower-order dependence on N_s can be considered. This might be the case for example where an antenna or single-source problem is being modeled or where the fields scattered from an object for a limited number of incidence angles or sources are needed. In such situations, iterative techniques provide a logical alternative to a direct solution. While the details can vary since a variety of iteration strategies are available, the basic goal is to achieve an acceptably accurate solution after a number of iterations N_i such that $N_i \ll N_s$. Since each iteration requires on the order of N_s^2 (or fewer) operations, this means that a solution for one RHS might be achieved after $N_s^2 N_i$ operations rather than the N_s^3 required by a direct solution.

Some of the better-known iteration techniques are Jacobi, Gauss-Seidel, and Gauss-Seidel with simultaneous over relaxation [Press, et. al. (1986)], with the first being of interest only in a comparative sense. They vary basically regarding how the updating is performed at each iteration step.

Another iteration procedure that has received much attention recently is the conjugate-gradient technique (CGT) [Sarkar and Rao (1984), Sarkar (1986)]. This approach develops a solution from a sequence of entire-domain components for the unknown, selected numerically such that they are "operator orthogonal". This means that taking the inner product of the j 'th component of the solution with the operator acting on the k 'th component produces the Kronecker delta function, δ_{jk} , thus permitting a solution to be developed term by term. These entire-domain components are necessarily constructed as linear combinations of the basis functions selected for discretizing the operator and computing the linear system of equations which replaces it numerically. In essence, the CGT transforms these "discretization" bases into those from which the eventual solution is constructed. Therefore, unless computation of the discretized operator is to be repeated during the iteration process, the form of each of the components in the solution is thereby determined by this initial basis-function selection.

An especially attractive property of CGT is that convergence is guaranteed after $N_i = N_s$ iterations. Unless convergence is achieved such that $N_i \ll N_s$, however, it should be noted that up to N_s^3 operations possibly could be needed. The convergence rate of CGT, as is the case of other iterative techniques, is determined by the eigenvalue spectrum of the operator matrix, with the best situation being all eigenvalues nearly equal. In this respect, CGT "prefers" a well-conditioned matrix, as one measure of matrix conditioning is a small ratio of the maximum to minimum eigen value. A distinct advantage of CGT is that a quantitative error measure is provided at each iteration step so that there is no numerical uncertainty about when the iterations can be concluded.

Iterative techniques have also been based on use of a Fourier transform, normally implemented via an FFT. The advantage of so doing is that integral equations of the kind shown in Table 6 have the form of a spatial convolution between the unknown and the kernel function. As is well known, the transform of a convolution becomes a product in the waveform or spectral domain, in this case the spatial-frequency domain. Consequently, the order of the discretized form of the integral equation is decreased from a squared dependence on the number of unknowns (the N_s^2 coefficients of the original impedance matrix) to a linear one (the product of two N_s 'th-sample spectra). But since the boundary condition(s) is expressed in real space and the product in spectral space, iteration to achieve a valid solution then becomes necessary. This basic idea has been described as the k -space transform [Bojarski (1972)], the spectral-iterative technique [Mittra and Kastner (1981), Kastner and Mittra (1983a, 1983b)], and the FFT approach [Sarkar, Arvas and Rao (1986), Sarkar (1986)].

The CGT has also been combined with a spectral approach using an FFT to achieve a further increase in computational efficiency [Sarkar (1986)]. For the special case of a straight wire for example, Sarkar obtains a solution time which increases only linearly with N_s , whereas solution via matrix factorization would increase as N_s^3 , and exploiting the Toeplitz symmetry would reduce this dependency to only N_s^2 (see table 12). Extensions of this basic idea to two and three dimensions (plates and cubes) have also been demonstrated [Sarkar (1986), Catedra, Cuevas and Nuno (1987)], where the computation time continues to grow linearly with N_s . Were this approach to be fully generalizable, it would imply that integral-equation modeling in the frequency domain of problems requiring one-, two- and three-dimensional sampling might be achieved in the limit for one right-hand-side in times which grow as f , f^2 and f^3 respectively. This represents a dramatic improvement over the results exhibited in Fig. 2 where solutions only using matrix factorization have been considered. We note that the minimum computation time that might be expected for any problem involving N_s samples should be expected to grow as N_s^p , where $p \geq 1$.

2. Exploiting Problem Symmetries

Problem symmetries provide one of the more effective means of reducing both computer time and storage.

The three basic types of symmetry are reflection, rotation, and translation symmetry. The first arises when an object is reflected about 1, 2 or 3 planes, as demonstrated by a straight wire, rectangular plate and right angle quadrilateral. Rotation (or circulant) symmetry occurs when an object reproduces itself upon rotation through $2\pi/n$ radians where the number of rotational sectors is denoted by the integer n . When n is finite, as for a regular polygon having n sides, the rotation symmetry is discrete. But n can also be effectively infinite, as for a circular loop which exhibits continuous rotation symmetry. The relationship between discrete and continuous rotational symmetry is equivalent to that between the discrete and continuous Fourier transform. Translation symmetry results when an object is created by rectilinear translation of some fixed shape, for which a straight wire again serves as the most basic example [Medgeys-Mitschang and Putnam (1983)].

Such symmetries can occur singly or in combination and can involve one or more unknowns per symmetric sector. The right circular cylinder provides an example that exhibits all three types (3 reflection planes, rotation about its axis, and translation along its axis). The computational benefits of symmetry accrue both in reducing computer storage and solution time because object symmetry creates a repetitive pattern in the direct matrix, the solution matrix, and the resulting solution, thus reducing the number of

needed operations. We further discuss each symmetry type briefly below. Exploiting object symmetry requires that the excitation (RHS) be decomposed in the same manner as the object itself, in a fashion analogous to how a plane wave is expanded in circular harmonics when incident on an infinite, circular cylinder.

a. Reflection symmetry

Consider the example of one-plane reflection symmetry provided by a single horizontal wire located over a perfectly conducting, infinite plane. In this case, the direct matrix takes the form

$$Z = \begin{bmatrix} Z_{ii} & Z_{is} \\ Z_{si} & Z_{ss} \end{bmatrix}$$

where the subscripts "i" and "s" refer to the image fields and the self fields respectively and we observe that $Z_{ii} = Z_{ss}$ and $Z_{si} = -Z_{is}$. But since the current I_s on the image will be the negative of the current I_i on the actual wire for any excitation applied to the latter, the direct matrix simplifies to

$$Z = Z_{ii} - Z_{is}$$

Similarly, if the wire were to be parallel to a perfect magnetic interface, then the image current would be equal to the actual current and the direct matrix simplifies to

$$Z = Z_{ii} + Z_{is}$$

These two cases correspond respectively to odd- and even-mode excitation. Compared with an unsymmetric problem having the same number of unknowns, the storage is reduced by one-half and the solution time by one-fourth. We note that this particular problem exhibits not only object and hence direct-matrix symmetry, but solution symmetry as well due to the constraints imposed by the image plane.

In general, the computer time for reflection symmetry can be expressed by

$$T_{\text{refl}} = A_{\text{fill}}(N_s/2^p)^2 + A_{\text{solve}}m(N_s/2^p)^3 \quad (4a)$$

where $p = 1, 2$, or 3 is the number of reflection planes, and $m = 1, \dots, 2^p$ is the number of excitation modes for which a solution is required.

b. Rotation Symmetry

In the case of n 'th-order rotation symmetry where the object has n discrete sectors (e.g., an n -sided polygon), the direct matrix has the general form

$$Z = \begin{bmatrix} Z_1 & Z_2 & \dots & Z_n \\ Z_n & Z_1 & \dots & Z_{n-1} \\ \vdots & \vdots & \ddots & \vdots \\ Z_2 & Z_3 & \dots & Z_1 \end{bmatrix}$$

where each Z_i can be a single coefficient or a matrix of order n_s where there are $n_s = N_s/n$ unknowns per rotational sector. For the circulant matrix which results from rotation symmetry and for any $k > n$, $Z_{ij} = Z_{i-k, j-k}$, where $\alpha, \beta = 1$ if $i+k > n, j+k > n$ respectively and are otherwise zero. Rotation symmetry thus produces a special Toeplitz matrix in that it is not only diagonal but has only n independent coefficients or blocks rather than the $2n-1$ of the usual Toeplitz form.

Because of its structure, a circulant matrix yields solutions that are given by a Fourier series of discrete, orthogonal modes.

Solution of the original matrix thus can be transformed to solution of $m = 1$ to n reduced or modal matrices, the actual number m needing solution depending on the kind of excitation being used. Each of these reduced matrices yields a solution for one of the m modes out of which the general, overall solution is constructed. Reduction of the original direct matrix into these modal matrices can occur because the angular variation of each solution mode is known. This is analogous to how knowledge that a horizontal wire over a perfectly conducting ground plane has oppositely directed currents is used in the case of reflection symmetry. In the case of rotation symmetry and mode number m for example, the solution in sector r is related to that in sector s by $\exp[i m(r-s)]$. Therefore, knowledge of the solution in sector $s=1$ is sufficient to construct the modal solution for any mode m .

The solution time for a problem having rotation symmetry is given for the general case by

$$T_{\text{rot}} = A_{\text{fill}}N_s^2/n + A_{\text{solve}}m(N_s/n)^3 \quad (4b)$$

where the fill time is thus reduced by a factor of n and the solution time by a factor of at least n^2 relative to a problem without symmetry. Similarly, the storage is reduced by a factor of n as well. It is worth noting that when $n = N_s$, which is the case of a single, n -sided polygon, the solution time is proportional to $15m \leq N_s$.

c. Translation Symmetry

A finite object having t sections of translation symmetry produces a matrix having the structure

$$Z = \begin{bmatrix} Z'_1 & Z'_2 & \dots & Z'_n \\ Z'_2 & Z'_1 & \dots & Z'_{n-1} \\ \vdots & \vdots & \ddots & \vdots \\ Z'_n & Z'_{n-1} & \dots & Z'_1 \end{bmatrix}$$

which is similar in appearance to a circulant matrix in that it is a Toeplitz matrix. Again, each Z_i can be a single coefficient or a matrix of order n_s with $n_s = N_s/t$ unknowns per t cross sections in the direction of translation. The primes indicate the possibility that "boundary terms" which arise at either end of the object may be slightly different from those in the interior, depending on the numerical treatment employed.

While the matrices produced by rotation and translation symmetry are both of Toeplitz form, they exhibit subtle though significant differences. For rotation symmetry, the inverse (solution) matrix is circulant as well, but for translation symmetry the inverse (solution) matrix is instead the sum of two products of two triangular Toeplitz matrices [Bitmead and Anderson (1980)]. One consequence of this difference is that the computer time needed for translation symmetry is not reduced as much as for rotation symmetry. In particular, rotation symmetry results in a solution time which varies between N_s^0 and N_s^1 , depending on the number of excitation modes, while that for translation symmetry can be as high as N_s^2 depending on the algorithm used. The overall computer time associated with translation symmetry is approximated by [Watson (1973)]

$$T_{\text{trans}} = A_{\text{fill}}N_s + A_{\text{solve}}n_s^2 t(\log_2 t)^2 \quad (4c)$$

where the $t(\log_2 t)$ term arises from transforming the Toeplitz matrix of order t into a circulant matrix of order $2t$ [Gustavson and Yun (1979), Bitmead and Anderson (1980)]. A slight further reduction is realized when the Toeplitz matrix is also banded [Jain (1978)].

d. Combinatorial Symmetry

When a problem possesses two or more kinds of symmetry it is possible to exploit their combination as a factored sequence to obtain a very significant further reduction in solution time. The possibility for doing so depends on the extent to which the modes which characterize each kind of symmetry remain orthogonal or separable. When mode coupling occurs between separately symmetric objects because they are not symmetrically oriented, as demonstrated by co-planar, identical, polygonal loops, then although the self-interaction of each object remains separable their interaction is not. If the loops are coaxial with their sides aligned however, then both the rotation and reflection symmetry can be fully exploited. One such example is given below to demonstrate the effects of combinatorial symmetry.

e. Example of Symmetry Application

Because the full impact of symmetry may be best appreciated by using a numerical example, let us briefly consider modeling a coaxial array of N_p identical, parallel, regular polygons having N_s sides so that $N_t = N_p N_s$. This problem leads to an overall direct matrix of N_t^2 th order which can be put in a block circulant array of $N_p \times N_p$ translation matrices each of order N_s . Since the sum of Toeplitz matrices is a Toeplitz matrix, the direct, block circulant matrix can be reduced to N_p Toeplitz matrices each of order N_s . Each of these N_s matrices can then be solved in a time proportional to N_s^3 .

For purposes of illustration, let $N_p = N_s = 1,000$ for a total of 1,000,000 unknowns, and assume that NEC [Numerical Electromagnetics Code, Burke and Poggio (1980)] is being run on a CDC-7600 computer for which $A_{\text{min}} = 1.4 \times 10^{-7}$ hour and $B_{\text{min}} = 5.6 \times 10^{-10}$ hour. The overall computer time can thus be concluded to change from (ignoring the fill time which is insignificant relative to the solution time)

$T = 5.6 \times 10^8$ hours (64,000 years) without symmetry;
to $T = 5.6 \times 10^2$ hours with rotation symmetry utilized;
to $T = 5.6 \times 10^{-1}$ hours (534 min) with both symmetries exploited.

A time reduction by a factor of $\sim N_s^{3/2} = 10^3$ is achieved in this case. Assuming that roundoff or other errors do not invalidate the numerical results for problems having this number of unknowns, the possibility of solving extremely large problems having high-order symmetry seems feasible. This is worth considering not only for those problems of practical interest that do have exploitable symmetries, but for testing other techniques for reducing computer time such as the near-neighbor approximation.

3. Near-neighbor Approximations

Motivation for the near-neighbor approximation (NNA) comes from the possibility that a window can be defined outside of which interactions are small enough to be ignored while still retaining acceptable accuracy in the solution. The NNA can be applied spatially in which case the rationale for ignoring interactions comes from the geometric attenuation of the fields with increasing distance from the source [Ferguson, et. al. (1976)]. A modal NNA can also be implemented for which the rationale is provided by the observation that coupling between modes tends to decrease as the difference between their mode numbers increases [Medgyesi-Mitschang and Putnam (1983), (1985)]. The benefit in either case is that a direct matrix is produced having a smaller proportion of nonzero coefficients as object size increases, with a consequent potential reduction in both computer storage and solution time.

If the average number of nonzero coefficients in each row of the sparse, direct matrix which results from the NNA is W , then the matrix fill time increases only as WN_s rather than N_s^2 . The subsequent solution time however depends on a number of factors, the most important one being the pattern of non-zeros in the direct matrix. For the one-dimensional problem of a straight wire for example, the NNA direct matrix is banded of width W about the main diagonal and can be solved without further approximation in a time proportional to $N_s W^2$ and a savings relative to a full-matrix solution $\sim N_s^2$.

For two- and three-dimensional problems, it is not generally possible to exploit matrix sparseness as easily and fully as a banded matrix having many zero coefficients within the band is produced. The result for a square plate and a penetrable cube is a solution time proportional to N_s^2 and $N_s^{7/2}$ respectively with a corresponding savings over a full-matrix solution of order N_s and $N_s^{1/2}$. While these may not be insignificant reductions in solution time for large-enough N_s , realizing the fullest benefits of the NNA requires some other approach to exploit matrix sparseness, e.g., an iterative procedure.

4. Other

Other numerical techniques for reducing computer-time requirements might be considered. Two we discuss here are "borrowed" from the area of signal processing, adaptive modeling for reducing the number of unknowns needed to achieve a desired accuracy, and model-based parameter estimation for increasing the utilization efficiency of what has been computed.

a. Adaptive Modeling

Most CEM seems to involve sampling of both the unknowns to be solved and the equations to be matched in a way that is determined more by problem geometry than electromagnetic requirements. Perhaps the simplest example is that of a straight wire for which equal-length segments are most often employed when using a sub-domain basis, resulting in equally spaced equation-match points as well when using collocation. This modeling strategy is used more for convenience than from knowledge that it provides the best accuracy for a given number of unknowns. The use of adaptive sampling could lead to improved modeling performance by adjusting the sampling to the requirements of a particular problem.

There are at least two ways by which adaptive sampling might be implemented. The simpler approach which we denote as static adaptation, would develop pre-determined sampling strategies for various problem classes which would then be used when problems from among them are being modeled. The other, more complex but potentially of better performance, would involve dynamic adaptation to increase the sampling during the course of the modeling computation. The latter approach would require quantitative assessment of the modeling error at a given stage in the computation together with a strategy for changing the sampling to improve the modeling accuracy. One version of the conjugate gradient technique is adaptive in this sense. A desirable attribute of dynamic adaptation is to utilize as fully as possible all previous computations when increasing the sampling at a given stage in the process, so that information already computed is not discarded.

b. Model-based parameter estimation

Identifying trends in data by sequentially connecting adjacent points with straight lines or using smoothed approximations thereto is an exercise familiar to any engineering student, a procedure typically described as curve fitting. The basic idea is to see whether curves which "fit" the data exhibit any sort of explainable behavior which reveals some underlying physics and/or mathematics. A classical example of curve fitting is polynomial interpolation and extrapolation as previously discussed in connection with selecting basis and testing functions for using the moment method.

When the curve-fitting procedure employs an analytical description based on the physics of the process which produced the data, the exercise may be more accurately described as one of model-based parameter estimation (MBPE). The analytical description provides the model whose parameters are adjustable constants needing numerical determination before the model is quantifiable. These constants are estimated by fitting the model to the data, usually in some least-mean-square sense. In electromagnetics, and in the context of our previous discussion that source-field relationships establish a transfer-function representation, we observe that MBPE is a tool for solving the inverse problem. One example is that of finding the complex resonances or poles of objects from their temporal or spectral responses using an exponential- or pole-series model.

Besides being a tool for treating inverse problems, MBPE also provides a means for developing more efficient tools for the direct or analysis problem. The principle in the latter case comes from the observation that the mathematical complexity of a formal, analytical solution can sometimes obscure the physical simplicity of what that solution describes. As a specific example we can cite the Sommerfeld solution of the interface problem whose treatment has drawn much attention. While the Sommerfeld integrals can be complicated to evaluate either analytically or numerically, the fields they describe exhibit reasonably simple spatial variation. It was this simplicity that was exploited by Brittingham et. al. (1977) in linearly interpolating in a spatial mesh of pre-computed Sommerfeld fields to develop the direct matrix for wires near an interface. By replacing direct Sommerfeld-integral evaluation with this simple curve-fitting procedure, the computer time required for using such models was reduced by a factor of 100 or more. Later, an analytically more accurate model with even better results than linear interpolation was developed based on asymptotic approximations to the Sommerfeld integrals [Burke and Miller (1984)].

Modeling interfaces appears to be just one of many direct problems in electromagnetics that are suited to MBPE for increasing their numerical efficiency. Other examples of MBPE include developing more efficient algorithms for computing the radiation pattern of a parabolic reflector [Bucci et. al. (1983)], and synthesizing and imaging aperture and antenna source distributions [Miller and Lager (1978), Inagake and Garbacz (1982), and Miller (1983)].

c. Numerical Green's function

The solution matrix that is obtained for a given object or structure in essence is the numerical equivalent of a generalized Green's function in that the boundary conditions prescribed in obtaining the matrix are satisfied over the structure's surface whatever the excitation to which it is exposed. Whether that excitation is caused by an incident plane wave to model its scattering properties, or a localized field to model its radiation characteristics, the solution matrix represents a self-contained numerical analog of the structure it approximates.

In many applications however, for example when antenna placement on the structure is being evaluated, the basic solution matrix is incomplete without including the antenna(s) in the model as well. Although their inclusion will not significantly increase the size of the direct matrix that must be solved when these added antennas are small relative to the overall structure, each new position they occupy changes that part of the direct matrix to which they contribute. In such cases it is computationally advantageous to partition the direct matrix into two parts, one for the basic structure and a remainder that accounts for the antenna-structure interaction. We thus have

$$Z = \begin{bmatrix} Z_{ss} & Z_{sa} \\ Z_{as} & Z_{aa} \end{bmatrix}$$

where "s" and "a" subscripts denote the structure and antenna respectively. A solution to the moment-method problem can then be expressed as

$$\begin{aligned} I_s &= Y_{ss} [E_s - Z_{sa} I_a] = I_s + I_s \\ I_a &= Y_{aa} [E_a - Z_{as} I_s] = I_a + I_a \end{aligned}$$

where I and E are the self- and mutual-interaction currents respectively. By substituting I_s into the second equation, we obtain

$$\begin{aligned} I_a &= (1 - Y_{aa} Z_{as} Z_{ss}^{-1} Z_{sa})^{-1} (I_a - Y_{aa} Z_{as} Y_{ss} E_s) \\ I_s &= I_s - Y_{ss}^{-1} Z_{sa} Z_{aa}^{-1} Z_{as} I_a \end{aligned}$$

so that a matrix of only order "a*a" rather than "s*s" needs to be solved for each new antenna location. The role played by Y_{ss} in this case is analogous to that of an analytic Green's function for various special geometries as discussed above. The reduction in computer time is of order "s" relative to solving the combined problem.

VII. VALIDATION, ERROR CHECKING, AND ERROR ANALYSIS

A. Modeling Uncertainties

The process of proceeding from an original physical problem to computed results is one which is subject to numerous uncertainties caused by a variety of factors. Perhaps foremost among these factors is the degree of arbitrariness associated with many of the choices that are made by the code developer and/or modeler in the course of eventually obtaining numerical results. Whereas the numerical evaluation of classical boundary-value problems such as scattering from a sphere is numerically robust in the sense that different workers using different computers and different software can obtain results in agreement to essentially as many significant figures as they wish, the same observation cannot be made for moment-method modeling.

Modeling uncertainties can be assigned to two basic error categories, a physical modeling error ϵ_p and a numerical modeling error ϵ_n as outlined in Table 18. The former is due to the fact that for most problems of practical interest varying degrees of approximation are needed in developing a simplified or idealized problem representation that will be compatible with the computer code to be used for the modeling computations. The latter is due to the fact that the numerical results obtained are almost invariably only approximate solutions to that idealized representation. We note that although an analytical expression may in principle represent a formally exact solution, the process of obtaining numerical results in that case is still one which inevitably involves finite-precision evaluation of the formal solution.

By its very nature, the physical modeling error requires some kind of measurement for its determination, except for those few problems whose analytical solution in principle involves no physical idealization nor subsequent numerical approximation. One example of such problems is that of determining the scattering or radiating properties of the perfectly conducting or dielectric sphere.

The numerical modeling error is itself comprised of two components in general, the determination of which would normally involve one or more kinds of computation. The first and generally more important of these components is the solution error which arises because the computer model used, even if solved exactly, would not provide an exact solution for the idealized problem representation. The solution error arises essentially due to the fact that the computer model is solved using a finite number of unknowns. The other, generally less important contributor to the numerical modeling error is the equation error which arises because the numerical results obtained from the computer model used may not numerically satisfy the modeling equations. The equation error may be caused both by roundoff due to the computer word size as well as the solution algorithm employed, as in the case of iteration, for example. The impact of equation error can be expected to increase with increasing condition number of the direct matrix. As an indicator of the kinds of issues that must be kept in mind when developing a numerical model, a brief list of generic modeling guidelines is presented in Table 19.

B. Validation and Error Checking

One of the most time consuming and long lasting of the tasks associated with any model development is that of validation. Long after work on the model has been completed questions will continue to arise about whether a given result is valid or whether the model can be applied to a given problem. There are essentially two kinds of validation procedures that can be considered to answer such questions:

- 1) Internal Validation, a check that can be made concerning solution validity within the model itself as summarized in Table 20.
- 2) External Validation, a check that utilizes information from other sources which could be analytical, experimental or numerical as summarized in Table 21.

Existing computer models often do not perform internal checks on the results they produce, but instead leave that as an exercise for the user. For example, NEC (Numerical Electromagnetics Code), one of the more widely used models, could provide and indeed has been exercised to give various kinds of checks relating to power balance, reciprocity and boundary-condition matching. But the software to do this is not an integral part of the code, generally being "patched in" by the user for a particular problem and check. It would be of extremely great potential value if a variety of such checks could be built into the code and exercised as desired by the modeler.

1. Internal Checks

As a particular example of the ways in which internal checks could be used, consider the case when a problem new to the modeler is being implemented and the initial results are obtained. Present practice usually involves "eye-balling" the data to see if it feels right, perhaps having first run some documented test cases to verify code performance. Since these test cases would not be likely to closely resemble the new problem, their successful solution would not provide much insight concerning the new results. If however, a series of checks built into the code could then be exercised at the modeler's discretion to verify that conditions necessary for a valid solution of Maxwell's Equations are satisfied, confidence in the model's reliability could be enhanced. These checks might range from being as exhaustive as boundary-condition matching would be, to being fairly simple, such as checking for reciprocity and power conservation. They could only be viewed as necessary but not sufficient conditions for solution validity, and could only involve such behavioral aspects as are not implicit in the model already (e.g., some formulations produce symmetric matrices so that bi-static scattering and transmit-receive reciprocity are assured). It would seem feasible to develop a figure-of-merit from the results of such checks that would provide in a single number a "quality factor" for the solution.

2. External Checks

The second kind of check involves the use of independent data from other sources. Perhaps the most convincing overall is experimental data, but analytical or numerical results should be comparably useful. Indeed, one of the most convenient computational checks would be provided by a code that permits two different models to be developed for the same problem, for example by incorporating user-selectable basis and weight functions. For greatest utility, such checks ideally should not be of single-point nature, for example to compare results for input impedance at a single frequency. Experience shows that computer models produce results that exhibit slight frequency shifts, angle shifts or spatial shifts in field quantities with respect to "exact" solutions, or even other computer models. Consequently, global comparisons are usually more meaningful, but even they may not be straightforward to implement and interpret. If the shifts mentioned are observed, it would seem appropriate to develop a correlation measure to establish the minimum squared difference between the two results as they are shifted along the axis of the common variable. For other models and applications, the results may be even less directly comparable, as is the case for IE and DE modeling approaches. Some work is needed in the general area of how results from two different representations of the same problem can be most meaningfully compared.

VIII. CONCLUDING REMARKS

In the preceding discussion, we have presented a selective survey of computational electromagnetics. Attention has been directed to radiation and scattering problems formulated as integral equations and solved using the Method of Moments. Beginning from the viewpoint of electromagnetics as a transfer-function process, we concluded that the basic problem is one of developing source-field relationships, or field propagators. Of the various ways by which these propagators might be expressed, we briefly discussed the Maxwell curl equations and Green's-function source integrals as providing the analytical basis for moment-method computer models. We then considered at more length some of the numerical issues involved in developing a computer model, including the idea of sampling functions used both to represent the unknowns to be solved for and to approximate the equations that they must satisfy. Some of the factors involved in choosing these sampling functions and their influence on the computational requirements were examined. Next, we discussed some ways of decreasing the needed computer time based on either analytical or numerical approaches. Some closing comments were directed to the important problem of validation, error checking and error analysis. Throughout our discussion, emphasis has been given to implementation issues involved in developing and using computer models as opposed to exploring analytical details.

IX. ACKNOWLEDGEMENT

The author greatly appreciates the careful and knowledgeable review of this paper by Dr. Arthur D. Yahjian who made numerous valuable suggestions for its improvement. Thanks are also due to the Dr. Ray J. King, Chairman of the Wave Standards Propagation Committee whose invitation motivated the paper, and to Transactions Associate Editor Dr. Robert J. Mailloux who provided guidance and encouragement during its preparation.

X. REFERENCES

- Abramowitz, M., and I. A. Stegun (1964), *Handbook of Mathematical Functions*, Applied Mathematics Series, Vol. 55, Washington, National Bureau of Standards.
- Acton, F. S. (1970), *Numerical Methods that Work*, Harper and Row, New York.
- Andreasson, M. G. (1965), Scattering from bodies of revolution, *IEEE Trans. Antennas and Propagat.*, AP-13, 303-310. Correction Sep 66, 659.
- Arden, B. W., and K. N. Astill (1970), *Numerical Algorithms: Origins and Applications*, Addison-Wesley Publishing Co, Reading, MA.
- Banos, A. (1966), *Dipole radiation in the presence of a conducting half-space*, Pergamon Press, New York.
- Barber, P. W., J. F. Owen, and R. K. Chang (1982), Resonant scattering for characterization of axisymmetric dielectric objects, *IEEE Trans. Antennas and Propagat.*, AP-28, 168-172.
- Bennett, C. L., and W. L. Weeks (1970), Transient scattering from conducting cylinders, *IEEE Trans. Antennas and Propagat.*, AP-18, 627-633.
- Bitmead, R. R., and B. D. O. Anderson (1980), Asymptotically fast solution of Toeplitz and related systems of linear equations, *Linear Algebra and its Applications*, 34, 103-115.
- Bojarski, N. N. (1972), K-Space formulation of the acoustic scattering problem", Proc. Acoustical Soc. of Am. Symposium, Miami, FL, 102.
- Bornholdt, J. M., and L. N. Medgyesi-Mitschang (1986), Mixed domain Galerkin expansions in electromagnetic problems, National Radio Science Meeting, Boulder, CO, 170.
- Brittingham, J. N., E. K. Miller, and J. T. Okada (1977), A bivariate interpolation approach for efficiently and accurately modeling antennas near a half space, *Electron. Lett.* 13, 690-691.
- Bucci, O. M., G. D'Elia, G. Franceschetti, and R. Pierri (1983), Efficient computation of the far field of parabolic reflectors by pseudo-sampling algorithm, *IEEE Trans. Antennas and Propagat.*, AP-31, 931-937.
- Burke, G. J., and A. J. Poggio (1980), Numerical Electromagnetics Code (NEC)—Method of Moments Parts I, II, and III, NOSC TD 116, Naval Ocean Systems Center, San Diego, CA, (revised).
- Burke, G. J., and E. K. Miller (1984), Modeling antennas near to and penetrating a lossy interface, *IEEE Trans. Antennas and Propagat.*, AP-32, 1040-1049.
- Burke, G. J., and E. K. Miller (1986), A comparison of several methods for evaluating the field of a source near an interface, The South African Institute of Electrical Engineers, Joint Symposium on Antennas & Propagation and Microwave Theory & Techniques, Pretoria, SA.
- Burnside, W. D., and P. H. Pathak (1980), A summary of hybrid solutions involving moment methods and GTD, in *Applications of the Method of Moments to Electromagnetic Fields*, B. J. Strait, ed., The SCEE Press, St. Cloud, Florida.
- Catedra, M. F., J. G. Cuevas and L. Nuno (1987), A scheme to analyze conducting problems of resonant size using the conjugate-gradient method and the fast Fourier transform, presented at NATO Advanced Study Institute on Modeling and Measurement in Analysis and Synthesis Problems, Il Ciocco, Italy, August, and to appear in proceedings published by Sijthoff & Noordhoff, the Netherlands.
- Chao, H. H., and B. J. Strait (1971), Radiation and scattering by configurations of bent wires with junctions, *IEEE Trans. Antennas and Propagat.*, AP-19, 701-702.
- Einzig, P. D., and L. B. Felsen (1983), Rigorous asymptotic analysis of transmission through a curved dielectric slab, *IEEE Trans. Antennas and Propagat.*, AP-31, 863-870.
- Ferguson, T. R., T. H. Lehman, and R. J. Balestri (1976), Efficient solution of large moments problems: Theory and small problem results (Succinct p.), *IEEE Trans. Antennas and Propagat.*, AP-24, 230-235.
- Glisson, A. W., and D. R. Wilton (1980), Simple and efficient numerical methods for problems of electromagnetic radiation and scattering from surfaces, *IEEE Trans. Antennas and Propagat.*, AP-28, 593-603.
- Gustavson, F. G. and D. Y. Y. Yun (1979), Fast algorithms for rational Hermite approximation and solution of Toeplitz systems, *IEEE Trans. Circuits and Systems*, IT-18, 725-730.
- Harrington, R. F. (1961), *Time-Harmonic Electromagnetic Fields*, McGraw Hill Book Company, New York.

- Harrington, R. F. (1968), *Field Computation by Moment Methods*, Macmillan, New York.
- Harrington, R. F., and J. R. Mautz (1975), An impedance sheet approximation for thin dielectric shells, *IEEE Trans. Antennas and Propagat.*, AP-23 531-534.
- Harrington, R. F. (1980), Origin and development of the method of moments for field computation, Chapter 1 in *Applications of the Method of Moments to Electromagnetic Fields*, B. J. Strait, ed., The SCEE Press, St. Cloud, FL.
- Harrington, R. F., D. R. Wilton, C. M. Butler, R. Mittra, and C. L. Bennett (1981), *Computational Methods in Electromagnetics*, The SCEE Press, St. Cloud, Florida.
- Inagaki, N. and R. J. Garbacz (1982), Eigenfunctions of composite Hermitian operators with application to discrete and continuous radiating systems, *IEEE Trans. Antennas and Propagat.*, AP-30, 575-575. Correction, Nov 82, 1268.
- Itoh, T. (1986), An overview on numerical techniques for modeling miniaturized passive components, *Ann. Telecommun.*, 41, 449-462.
- Jain, A. K. (1978), Fast inversion of banded Toeplitz matrices by circular decompositions, *IEEE Trans. Acoustics, Speech, and Sig. Proc.*, ASSP-26, 121-125.
- Jeng, S.-K., and C.-H. Chen (1984), On variational electromagnetics: theory and application, *IEEE Trans. Antennas Propagat.*, AP-32, 902-907.
- Kastner, R. and R. Mittra (1983a), A spectral-iteration technique for analyzing scattering from arbitrary bodies—I: Cylindrical scatters with E-wave incidence, *IEEE Trans. Antennas Propagat.*, AP-31, 499-506.
- Kastner, R. and R. Mittra (1983b), A spectral-iteration technique for analyzing scattering from arbitrary bodies—II: Cylindrical scatters with H-wave incidence, (Commun.), *IEEE Trans. Antennas Propagat.*, AP-31, 535-537.
- Komiyama, A. (1984), On an arrangement of sources in the charge simulation method—case of E wave scattering by one perfectly conducting cylinder, *Trans. IECE of Japan*, J67-B, 953-959.
- Leontovich, M. A. (1948), *Investigations on Radiowave Propagation, Part II*. Moscow: Academy of Sciences.
- Leviatan, Y. and A. Boag (1987), Analysis of electromagnetic scattering from dielectric cylinders using a multifilament current model, *IEEE Trans. Antennas and Propagat.*, AP-35, 1119-1127.
- Lindell, I. V., E. Alanen, and K. Mannersalo (1985), Exact image method for impedance computation of antennas above the ground, *IEEE Trans. Antennas and Propagat.*, AP-33, 937-945.
- Lucke, W. S. (1951), Electric dipoles in the presence of elliptic and circular cylinders, *J. Appl. Physics*, 22, 14-19.
- Ludwig, A. C. (1986), A comparison of spherical wave boundary value matching versus integral equation scattering solutions for a perfectly conducting body, *IEEE Trans. Antennas and Propagat.*, AP-34, 857-865.
- Lytile, R. J. (1971), Far-field patterns of point sources operated in the presence of dielectric circular cylinders, *IEEE Trans. Antennas and Propagat.*, AP-19, 618-621.
- Mason, J. L., and W. J. Anderson (1985), Finite element solution for electromagnetic scattering from two-dimensional bodies, *Int. J. for Numerical Method in Eng.*, 21, 909-928.
- Maue, A. W. (1949), The formulation of a general diffraction problem by an integral equation, *Zeitschrift fur Physik*, 126, 601-618.
- Mautz, J. R., and R. F. Harrington (1969), Radiation and scattering from bodies of revolution, *Appl. Sci. Res.*, 20, 405-435.
- Mautz, J. R., and R. F. Harrington (1978), H-Field, E-Field, and combined-field solutions for conducting bodies of revolution, *A. E. U. (Germany)*, 32, 157-164.
- Medgyesi-Mitschang, L. N., and J. M. Putnam (1983), Scattering from finite bodies of translation: plates, curved surfaces, and non-circular cylinders, *IEEE Trans. Antennas and Propagat.*, AP-31, 847-852.
- Medgyesi-Mitschang, L. N., and J. M. Putnam (1985), Electromagnetic scattering from extended wires and two- and three-dimensional surfaces, *IEEE Trans. Antennas and Propagat.*, AP-33, 1090-1100.
- Mei, K. K. (1965), On the integral equations of thin wire antennas, *IEEE Trans. Antennas and Propagat.*, AP-13, 374-378.
- Mei, K. K. (1974), Unimoment method of solving antenna and scattering problems, *IEEE Trans. Antennas and Propagat.*, AP-22, 760-766.
- Mieras, H., and C. L. Bennett (1982), Space-time integral equation approach to dielectric targets, *IEEE Trans. Antennas and Propagat.*, AP-30, 2-9.
- Miller, E. K., and A. Olte (1966), Excitation of surface currents on a plasma-immersed cylinder by electromagnetic and electrokinetic waves. II. The inhomogeneous sheath, *Radio Science*, 1, 1425-1433.
- Miller, E. K., G. J. Burke, and E. S. Selden (1971), Accuracy-modeling guidelines for integral-equation evaluation of thin-wire scattering structures, *IEEE Trans. Antennas and Propagat.*, AP-19, 534-536.

- Miller, E. K., A. J. Poggio, G. J. Burke, and E. S. Selden (1972a), Analysis of wire antennas in the presence of a conducting half-space. Part I. The vertical antenna in free space, *Canadian J. of Physics*, 50, 879-888.
- Miller, E. K., A. J. Poggio, G. J. Burke, and E. S. Selden (1972b), Analysis of wire antennas in the presence of a conducting half-space. Part II. The horizontal antenna in free space, *Canadian J. of Physics*, 50, 2614-2627.
- Miller, E. K., and F. J. Deadrick (1975), Some computational aspects of thin-wire modeling, Chapter 4 in *Numerical and Asymptotic Techniques in Electromagnetics*, R. Mittra, ed., Springer Verlag, New York, 89-127.
- Miller, E. K., and D. L. Lager (1978), Radiation field analysis and synthesis using Prony's Method, *Electronics Letter*, 14, 180.
- Miller, E. K. (1981), Wires and wire grid models for radiation and scattering including ground effects, in *Theoretical Methods for Determining the Interaction of Electromagnetic Waves with Structures*, J. K. Skwirzynski ed., Sijthoff & Noordhoff.
- Miller, E. K. (1983), Imaging of linear source distributions, *Electromagnetics*, 3, 21-40.
- Mitzner, K. M. (1967), An integral equation approach to scattering from a body of finite conductivity, *Radio Science*, 2, 1459-1470.
- Mittra, R. (1973), ed., *Computer Techniques for Electromagnetics*, Pergamon Press, New York.
- Mittra, R. (1975), ed., *Numerical and Asymptotic Techniques in Electromagnetics*, Springer Verlag, New York.
- Mittra, R. and R. Kastner (1981), A spectral domain approach for computing the radiation characteristics of a leaky-wave antenna for millimeter waves (Commun.), *IEEE Trans. Antennas and Propagat.* AP-29, 652-654.
- Moore, J. and R. Pizer (1984), "Moment Methods in Electromagnetics: Techniques and Applications", Wiley & Sons, New York.
- Ney, M. M. (1985), Method of moments as applied to electromagnetic problems, *IEEE Trans. Microwave Theory Tech.*, MTT-33, 972-980.
- Oshiro, F. K. (1965), Source distribution techniques for the solution of general electromagnetic scattering problems, *Proc. First GISAT Symposium, Mitre Corp.*, vol. I, 83-107.
- Oshiro, F. K. and Mitzner (1967), Digital computer solution of three-dimensional scattering problems, *IEEE International Antennas and Propagation Symposium*, U. of Michigan, Ann Arbor, October, 257-63.
- Pathak, P. H. (1983), On the eigenfunction expansion of electromagnetic dyadic Green's functions, *IEEE Trans. Antennas and Propagat.*, AP-32, 837-846.
- Perine, J., and D. J. Buchanan (1982), Assessment of MoM Techniques for shipboard applications, *IEEE Trans. Electromagn. Comp.*, EMC-24, 32-39.
- Poggio, A. J., and E. K. Miller (1973), Integral equation solutions of three-dimensional scattering problems, Chapter 4 in *Computer Techniques for Electromagnetics*, R. Mittra ed., Pergamon Press, New York.
- Poggio, A. J., and E. K. Miller (1987), Low frequency analytical and numerical methods for antennas, in *Antenna Handbook*, V. T. Lo and S. W. Lee, ed., Van Nostrand Reinhold, New York.
- Popovic, B. D., M. B. Dragovic, and A. R. Djordjevic (1982), "Analysis and Synthesis of Wire Antennas", Research Studies Press, Letchworth, Hertfordshire, England.
- Press, W. H., B. R. Flannery, S. A. Teukolsky, and W. T. Vetterling (1986), *Numerical Recipes*, Cambridge University Press, London.
- Richmond, J. H. (1965), Digital computer solutions of the rigorous equations for scattering problems, *Proc. IEEE*, 53, 796-804.
- Richmond, J. H. (1966), A wire-grid model for scattering by conducting bodies, *IEEE Trans. Antennas and Propagat.*, AP-14, 782-786.
- Richmond, J. H., (1974), Radiation and scattering by thin-wire structures in a homogeneous, conducting medium, *IEEE Trans. Antennas and Propagat.*, AP-22, 365.
- Rytov, S. M. (1940), *J. Exp. Theor. Phys. USSR*, vol. 10, no. 2, p. 180 (in Russian).
- Sarkar, T. K. and S. M. Rao (1984), The application of the conjugate gradient method for the solution of electromagnetic scattering from arbitrarily oriented wire antennas", *IEEE Trans. Antennas Propagat.*, AP-32, no. 4, 398-403.
- Sarkar, T. K. (1986), The conjugate-gradient technique as applied to electromagnetic field problems, *IEEE AP-S Newsletter*, 28, no. 4, 5-14.
- Sarkar, T. K., E. I. Eruvas and S. M. Rao (1986), Application of FFT and conjugate gradient method for the solution of electromagnetic radiation from electrically large and small conducting bodies", *IEEE Trans. Antennas Propagat.*, AP-33, 635-640.
- Schaubert, D. H., D. R. Wilton, and A. W. Glisson (1984), A tetrahedral modeling method for electromagnetic scattering by arbitrarily shaped inhomogeneous dielectric bodies, *IEEE Trans. Antennas Propagat.*, AP-32, 77-85.
- Senior, T. B. A. (1960), Impedance boundary conditions for imperfectly conducting surfaces, *Appl. Sci. Res.*, vol 8(B), pp. 418-436.

- Senior, T. B. A. (1981), Approximate boundary conditions (Mini-rev.), *IEEE Trans. Antennas and Propagat.*, **AP-29**, 826-829.
- Strait, B. J., editor (1980), *Applications of the Method of Moments to Electromagnetic Fields*, the SCEEE Press, St. Cloud, Florida.
- Strait, B. J., and A. T. Adams (1980), On contributions at Syracuse University to the Moment Method, *IEEE Trans. Electromagn. Comp.*, **EMC-22**, 228-237.
- Stratton, J. A. (1941), *Electromagnetic Theory*, McGraw-Hill, New York.
- Stutzman, W. L. and G. A. Thiele (1981), *Antenna Theory and Design*, John Wiley and Sons, New York.
- Taflov, A., and K. Umashankar (1980), A hybrid moment method/finite difference time domain approach to electromagnetic coupling and aperture penetration in complex geometries, in *Applications of the Method of Moments to Electromagnetic Fields*, B. J. Strait, ed., The SCEEE Press, St. Cloud, Florida.
- Tesche, F. M., and A. R. Neureuther (1970), Radiation patterns for two monopoles on a perfectly conducting sphere, *IEEE Trans. Antennas and Propagat.*, **AP-18**, 692-694.
- Tesche, F. M. (1972), On the behavior of thin-wire antennas and scatterers arbitrarily located within a parallel-plate region, *IEEE Trans. Antennas and Propagat.*, **AP-20**, 482-486.
- Thiele, G. A., and T. H. Newhouse (1969), A hybrid technique for combining moment methods with the geometrical theory of diffraction, *IEEE Trans. Antennas and Propagat.*, **AP-17**, 62-69.
- Uslenghi, P. L. E. (1980), ed., *Electromagnetic Scattering*, Academic Press.
- Van Nostrand Scientific Encyclopedia (1976), Fifth Edition, Electromagnetic Phenomena, 919-923.
- Varadan, V. V. and V. K. Varadan (1982), T-matrix approach to acoustic, electromagnetic and elastic wave scattering, presented at *Research Techniques in Wave Propagation and Scattering*, Ohio State University, October.
- Wacker, P. F. (1981), Unified theory of near-field analysis and measurement: Scattering and inverse scattering, *IEEE Trans. Antennas and Propagat.*, **AP-29**, 342-351.
- Waterman, P. C. (1965), Matrix formulation of electromagnetic scattering, *Proc. IEEE* **53**, 805-12.
- Waterman, P. C. (1973), Numerical solution of electromagnetic scattering problems, in *Computer Techniques for Electromagnetics*, R. Mittra ed., Pergamon Press, New York.
- Watson, G. A. (1973), An algorithm for the inversion of block matrices of Toeplitz form, *J. of the Assoc. for Computing Machinery*, **20**, 409-415.
- Wilton, D. R., and C. M. Butler (1976), Efficient numerical techniques for solving Pocklington's equation and their relationships to other methods, *IEEE Trans. Antennas and Propagat.*, **AP-24**, 83-86.
- Wilton, D. R., and C. M. Butler (1981), Effective methods for solving integral and integro-differential equations, *Electromagnetics*, **1**.
- Wu, D. T., and D. C. Chang (1986), Current distribution on a thin wire in a multimode rectangular cavity, National Radio Science Meeting, University of Colorado, Boulder, CO.

Table 1
CLASSIFICATION OF MODEL TYPES IN CEM

FIELD PROPAGATOR Integral operator Differential operator Modal expansions Optical description	DESCRIPTION BASED ON: Green's function for infinite medium or special boundaries. Maxwell Curl Equations or their integral counterparts. Solutions of Maxwell's Equations in particular coordinate system and expansion. Rays and diffraction coefficients.
APPLICATION Radiation Propagation Scattering	REQUIRES: Determining the originating sources of a field. Obtaining the fields distant from a known source. Determining the perturbing effects of medium inhomogeneities.
PROBLEM TYPE Solution domain Solution space Dimensionality Electrical properties of medium and/or boundary Boundary geometry	CHARACTERIZED BY: Time or frequency. Configuration or wavenumber. 1D, 2D, 3D. Dielectric; lossy; perfectly conducting; anisotropic; inhomogeneous; nonlinear. Linear; curved; segmented; compound; arbitrary

Table 2
STEPS IN DEVELOPING A COMPUTER MODEL

STEP Conceptualization	ACTIVITY Encapsulating observation and analysis in terms of elementary physical principles and their mathematical description.
Formulation	"Fleshing out" of the elementary description into a more complete, formally solved, mathematical representation.
Numerical implementation	Transforming into a computer algorithm using various numerical techniques.
Computation	Obtaining quantitative results.
Validation	Determining the numerical and physical credibility of the computed results.

Table 3
DESIRABLE ATTRIBUTES IN A COMPUTER MODEL

ATTRIBUTE Accuracy	DESCRIPTION The quantitative degree to which the computed results conform to the mathematical and physical reality being modeled. Accuracy, preferably of known and better yet, selectable, value is the single most important model attribute. It is determined by the physical modeling error (ϵ_p) and numerical modeling error (ϵ_n) [see Section VII].
Efficiency	The relative cost of obtaining the <u>needed</u> results. It is determined by the <u>human</u> effort required to develop the computer input and interpret the output, and by the associated <u>computer</u> cost of running the model.
Utility	The applicability of the computer model in terms of problem size and complexity. Utility also relates to ease of use, reliability of results obtained, etc.

Table 4
 REPRESENTATIVE APPROXIMATIONS THAT ARISE IN MODEL DEVELOPMENT

APPROXIMATION	IMPLEMENTATION/IMPLICATIONS
Conceptualization	
Physical Optics	Surface sources given by tangential components of incident field, with fields subsequently propagated via a Green's function. Best for backscatter and main-lobe region of reflector antennas, from resonance region ($ka > 1$) and up in frequency.
Physical Theory of diffraction	Combines aspects of physical optics and geometrical theory of diffraction, primarily via use of edge-current corrections to utilize best features of each.
Geometrical Theory of Diffraction	Fields propagated via a divergence factor with amplitude of obtained from diffraction coefficient. Generally applicable for $ka > 2.5$. Can involve complicated ray tracing.
Geometrical Optics	Ray tracing without diffraction. Improves with increasing frequency.
Compensation Theorem	Solution obtained in terms of perturbation from a reference, known solution.
Born-Rytov	Approach used for low-contrast, penetrable objects where sources are estimated from incident field.
Rayleigh	Fields at surface of object represented in terms of only outward propagating components in a modal expansion.
Formulation	
Surface Impedance	Reduces number of field quantities by assuming an impedance relation between tangential E and H at surface of penetrable object. May be used in connection with physical optics.
Thin-Wire	Reduces surface integral on thin, wirelike object to a line integral by ignoring circumferential current and circumferential variation of longitudinal current which is represented as a filament. Generally limited to $ka < 1$ where a is the wire radius.
Numerical Implementation	
$\partial f / \partial x \rightarrow (f_i - f_j) / (x_i - x_j)$ $\int f(x) dx \rightarrow \sum f(x_i) \Delta x_i$	Differentiation and integration of continuous functions represented in terms of analytic operations on sampled approximations, for which polynomial or trigonometric functions are often used. Inherently a discretizing operation, for which typically $\Delta x < \lambda / 2\pi$ for acceptable accuracy.
Computation	
Deviation of numerical model from physical reality	Affects solution accuracy and reliability to physical problem in ways that are difficult to predict and quantify.
Non-converged solution	Discretized solutions usually converge globally in proportion to $\exp(-AN_p)$ with A determined by the problem. At least two solutions using different numbers of unknowns N_p are needed to estimate A [Miller, Burke, and Selden (1971)].

Table 5
COMPARISON OF STEPS IN TIME- AND FREQUENCY-DOMAIN MODELING

Frequency Domain $\partial/\partial t = i\omega$	(Maxwell's equations)	Time Domain Time dependent
$L(\omega, r-r')f(\omega, r') = g(\omega, r)$	(Plus boundary conditions, etc.)	$L(r-r')f(\tau, r') = g(t, r)$ $\tau = t - r-r' /c$
where r is observation point, r' is source point, c is speed of light in medium		
$L(\omega, r-r')$ depends on ω and $ r-r' $	(Note that)	$L(r-r')$ depends on $ r-r' $ only
$\sum_{i,j=1}^{N_s} Z_{ij} f_j = g_i$ where N_s is number of space samples	(Apply MoM to get Nth order system)	$\sum_{i,j=1}^{N_t} Z_{ij} f_{jk} = g_{ik}$ $i,j = 1, \dots, N_t; k = 1, \dots, N_s$ where N_t is number of time samples
$f_i = \sum_j Y_{ij} g_j$ where $Y = Z^{-1}$	(Matrix manipulation yields)	$f_{ik} = \sum_j Y_{ij} g_{jk}$ where $Y = Z^{-1}$ (Inversion needed only once and only for implicit solution)
Solution obtained for many sources but single frequency	(Observe)	Solution obtained for single source but many frequencies (using Fourier transform)
Do for $m = 1, \dots, N_f$ frequencies to get f_{im}	f_{im} and f_{ik} are related by Fourier transform	—

Table 6
SOME INTEGRAL EQUATIONS WIDELY USED IN ELECTROMAGNETICS

FREQUENCY DOMAIN	COMMENTS
Magnetic-Field Integral Equation $\oint_S J_s(r) \cdot (1/2\pi) \nabla \times \int_S f(r') \nabla g(r, r') da' = 2\pi \nabla H^{inc}(r)$ $= L_H(J)$	Suited for smooth closed bodies. Note source term is physical optics current [Maue (1949), Oshiro (1965), Oshiro and Mitzner (1967)].
Electric-Field Integral Equation $(1/2\pi) \nabla \times \int_S [i\omega \mu J_s(r')] \nabla g(r, r') - [n \cdot E(r')] \nabla g(r, r') da' = 2\pi \nabla E^{inc}(r)$ $= L_E(J)$	Applicable to general bodies such as wires, plates, and shells. [Maue (1949), Andreason (1965)].
Combined-Field Integral Equation (CFIE) $L_H(J) + \alpha L_E(J)/\eta = 2\pi \nabla H^{inc}(r) + 2\alpha \nabla E^{inc}(r)/\eta$	Eliminates the spurious solutions that occur in either the EFIE or MFIE at interior resonances [Mitzner (1967), Mautz and Harrington (1978)].
TIME DOMAIN	
Magnetic-Field Integral Equation $J_s(r, t) - (1/2\pi) \nabla \times \int_S [1/R + c^{-1} \partial/\partial t'] [J_s(r', t')] \times R / R^3 ds' = 2\pi \nabla H^{inc}(r, t)$	Limitations similar to FD MFIE [Bennett and Weeks (1970)].
Electric-Field Integral Equation $(1/2\pi) \nabla \times \int_S [-\partial J_s(r', t')/\partial t' + [1/R + c^{-1} \partial/\partial t'] \times [n' \cdot E(r', t') R / R]] / R ds' = 2\pi \nabla E^{inc}(r, t)$	Applicability comparable to FD EFIE [Mieras and Bennett (1982)].
where: r and t are the observation position and time; r' and t' are the source position and time; $t' = t - R/c$ is the retarded time; $R = r-r' $ with r and r' on the surface S ; $g(r, r') = e^{-ikR}/R$	
\oint denotes a principle-value integral n, t are unit normal and tangent vectors to the surface α is an adjustable parameter to determine relative weighting of MFIE and EFIE contributions, with best results occurring for $0.25 \leq \alpha \leq 1$ η is free-space impedance J_s is induced surface current $\nabla \times H$	

Table 7
COMPARISON OF IE- AND DE-FIELD PROPAGATORS
AND THEIR NUMERICAL TREATMENT

	DIFFERENTIAL FORM	INTEGRAL FORM
Field Propagator	Maxwell Curl Equations	Green's function
Boundary Treatment		
At infinity (radiation condition)	Local or global "lookback" to approximate outward propagating wave	Green's function
On object	Appropriate field values specified on mesh boundaries to obtain stair-step, piecewise linear or other approximation to the boundary	Appropriate field values specified on object contour which can in principal be a general, curvilinear surface, although this possibility seems to be seldom used
Sampling Requirements		
No. of space samples	$N_s \propto (L/\Delta L)^D$	$N_s \propto (L/\Delta L)^{D-1}$
No. of time steps	$N_t \propto (L/\Delta L) \approx cT/\delta t$	$N_t \propto (L/\Delta L) \approx cT/\delta t$
No. of excitations (right-hand sides)	$N_{rh} \propto (L/\Delta L)$	$N_{rh} \propto (L/\Delta L)$
Linear system	Sparse, but larger	Dense, but smaller. In this comparison, note that we assume the IE permits a sampling of order one less than the problem dimension, i.e., inhomogeneous problems are excluded.
L is problem size		
D is no. of problem dimensions (1,2,3)		
T is observation time		
ΔL is spatial resolution		
δt is time resolution		
Dependence of Solution Time on Highest-Order Term in $(L/\Delta L)$		
Frequency domain	$T_s \propto N_s^{2(D-1)/D+1}$ $= (L/\Delta L)^{2D-1}$	$T_s \propto N_s^3 = (L/\Delta L)^{3(D-1)}$
Time domain		
Explicit	$T_i \propto N_s N_t N_{rh}$ $= (L/\Delta L)^{D+1+r}$	$T_i \propto N_s^2 N_t N_{rh}$ $= (L/\Delta L)^{2D+1+r}; 0 \leq r \leq 1$
Implicit	$T_i \propto N_s^{2(D-1)/D+1}$ $= (L/\Delta L)^{2D-1}, D=2,3;$ $\propto N_s N_t N_{rh} = (L/\Delta L)^{2+r}, D=1; 0 \leq r \leq 1$	$T_i \propto N_s^3 = (L/\Delta L)^{3(D-1)}$

Comments Note that D is the number of spatial dimensions in the problem, and is not necessarily the sampling dimensionality d. The distinction is important because when an appropriate Green's function is available, the source integrals are usually one dimension less than the problem dimension, i.e., $d = D-1$. An exception is an inhomogeneous, penetrable body where $d = D$ when using an integral equation. We also assume for simplicity that matrix solution is achieved via factorization rather than iteration, but that banded matrices are exploited for the DE approach where feasible. The solution-time dependencies given can thus be regarded as upper-bound estimates. See Table 12 for further discussion of linear-system solutions.

Table 8
RELATIVE APPLICABILITY OF IE- AND DE-BASED COMPUTER MODELS

TIME DOMAIN		ISSUE	FREQUENCY DOMAIN	
DE	IE		DE	IE
MEDIUM				
✓	✓	Linear	✓	✓
~	x	Dispersive	✓	✓
✓	x	Lossy	✓	✓
✓	~	Anisotropic	✓	✓
✓	x	Inhomogeneous	✓	x
✓	x	Nonlinear	x	x
✓	x	Time-varying	x	x
OBJECT				
~	✓	Wire	~	✓
✓	✓	Closed Surface	✓	✓
✓	✓	Penetrable Volume	✓	✓
~	✓	Open Surface	~	✓
BOUNDARY CONDITIONS				
✓	✓	Interior Problem	✓	✓
~	✓	Exterior Problem	~	✓
✓	✓	Linear	✓	✓
✓	✓	Nonlinear	x	x
✓	✓	Time-varying	x	x
~	x	Half-space	~	✓
OTHER ASPECTS				
~	~	Symmetry Exploitation	✓	✓
~	✓	Far-field Evaluation	~	✓
x	~	Number of Unknowns	~	✓
✓	~	Length of Code	~	x
SUITABILITY FOR HYBRIDING WITH OTHER:				
~	✓	Numerical Procedures	✓	✓
x	~	Analytical Procedures	~	✓
x	~	GTD	x	✓

where:

- ✓ signifies highly suited or most advantageous
- ~ signifies moderately suited or neutral
- x signifies unsuited or least advantageous

Table 9
SAMPLING OPERATIONS INVOLVED IN MoM MODELING

EQUATION	DE MODEL $L(s')f(s') = g(s')$	IE MODEL $L(s,s')f(s') = g(s)$
Sampling of: Unknown via basis- functions $b_j(s')$ using $f(s') = \sum a_j b_j(s')$.	Sub-domain bases usually of low order are used. Known as finite-difference procedure when pulse basis is used, and as finite-element approach when bases are linear.	Can use either sub-domain or entire-domain bases. Use of latter is generally confined to bodies of rotation. Former is usually of low order, with piece-wise linear or sinusoidal being the maximum variation employed.
Equation via weight functions $w_i(s)$ $\langle w_i(s), L(s,s') \sum a_j b_j(s') \rangle$ $= \langle w_i(s), g(s) \rangle$ to get $Z_{ij} a_j = g_i$.	Point-wise matching is commonly employed, using a delta-function. Pulse and linear matching are also used.	Point-wise matching is commonly employed, using a delta-function. For wires, pulse, linear, and sinusoidal testing is also used. Linear and sinusoidal testing is also used for surfaces.
Operator.	Operator sampling for DE models is entwined with sampling the unknown in terms of the difference operators employed.	Sampling needed depends on the nature of the integral operator $L(s,s')$. An important consideration whenever the field integral cannot be evaluated in closed form.
Solution of: $Z_{ij} a_j = g_i$ for the a_j .	Interaction matrix is sparse. Time-domain approach may be explicit or implicit. In frequency domain, banded-matrix technique usually employed.	Interaction matrix is full. Solution via factorization or iteration.

Table 10
EXAMPLES OF GENERIC BASIS/WEIGHT-FUNCTION COMBINATIONS

Method	jth Term of Basis	ith Term of Weight
Galerkin	$a_j b_j(r')$	$w_i(r) = b_i(r)$
Least square	$a_j b_j(r')$	$Q(r) \partial \epsilon(r) / \partial a_j$
Point matching	$a_j \delta(r-r')$	$\delta(r-r')$
General collocation	$a_j b_j(r')$	$\delta(r-r')$
Subsectional collocation	$U(r_j) \sum a_{jk} b_k(r')$	$\delta(r-r')$
Subsectional Galerkin	$U(r_j) \sum a_{jk} b_k(r')$	$U(r_i) \sum b_i(r)$

where:

r' and r denote source and observation points respectively

a_j, a_k are unknown constants associated with the j 'th basis function (entire domain) or the k 'th basis function of the j 'th subsection (sub-domain)

$U(r_j)$ is the unit sampling function which equals one on the k 'th sub-domain and is zero elsewhere

$b_j(r')$ is the j 'th basis function

$w_i(r)$ is the i 'th testing function

$\delta(r-r')$ is the Dirac delta function

$Q(r)$ is a positive-definite function of position

$\epsilon(r)$ is the residual or equation error

Table 11
EXAMPLES OF SPECIFIC BASIS/WEIGHT-FUNCTION
COMBINATIONS

Application	jth Term of Basis	ith Term of Weight
1D/Wires [Richmond (1965), (1966)]	Constant— $a_j U_j(s)$	Delta function— $\delta(s-s_j)$
1D/Wires [Chao and Strait (1971)]	Piecewise linear— $a_{j1}(s-s_j-\delta_j/2) + a_{j2}(s-s_j+\delta_j/2)$	Piecewise linear— $(s-s_j-\delta_j/2) + (s-s_j+\delta_j/2)$
1D/Wires [Mei (1965)]	3-term sinusoidal— $a_{j1} + a_{j2}\sin[k(s-s_j)] + a_{j3}\cos[k(s-s_j)]$	Delta function— $\delta(s-s_j)$
1D/Wires [Richmond 1974)]	Piecewise sinusoidal— $a_{j1}\sin[k(s-s_j-\delta_j/2)] + a_{j2}\sin[k(s-s_j+\delta_j/2)]$	Piecewise sinusoidal— $\sin[k(s-s_j-\delta_j/2)] + \sin[k(s-s_j+\delta_j/2)]$
2D/Surfaces [Oshiro(1965)]	Weighted Delta function— $a_j \delta(s-s_j) \Delta_j$	Delta function— $\delta(s-s_j)$
2D/ Rotational Surfaces [Mautz & Harrington(1969)]	Piecewise linear axially, and $\exp(in\theta)$ azimuthally	Same (Galerkin's method)
2D/Surfaces [Glisson & Wilton (1980)]	Piecewise linear	Same (Galerkin's method)
2D/Surfaces [Bornholdt and Medgeysi-Mitschang (1986)]	Piecewise linear subdomain/Fourier series entire domain	Same (Galerkin's method)
3D/Volumes [Schaubert et. al. (1984)]	Piecewise linear	Same (Galerkin's method)

Table 12
SUMMARY OF OPERATION COUNT FOR SOLUTION OF
GENERAL DIRECT MATRIX HAVING N_x UNKNOWN

Method	To Obtain Solution Matrix	To Obtain Solution	Comments
Cramer's Rule	Expand in co-factors leading to \rightarrow	$\sim N_x!$	Not an advisable procedure, but useful to illustrate just how bad the problem could be!
Inversion	N_x^3	N_x^3	Provides RHS-independent solution matrix
Factorization	$N_x^3/3$	N_x^3	RHS independent solution matrix
Iteration General	—	$N_x^2 \cdot N_x^2$	Each RHS requires separate solution
With FFT	—	$N_x \cdot N_x^2$	Same, plus applicability to arbitrary problems uncertain
Symmetry			
Reflection	$(1+2p) \times (N_x/2p)^3$	$N_x^3/2p$	For $p=1$ to 3 reflection planes
Translation (Toeplitz)	$n_x^3 [t(\log_e t)^2]$	N_x^3	For n_x unknowns per t sections of translation
Rotation (Circulant)	$\text{Log}_2(N_x) N_x^3/n^2$	N_x	For n rotation sectors and a complete solution
	$m \text{Log}_2(N_x)(N_x/n)^3$	m	For $m = 1$ to n modes
Banded General	$N_x W^2$	$N_x W$	For a bandwidth of W coefficients
Toeplitz	$N_x \text{Log}_2 N_x$	W^2	For a bandwidth of W coefficients

Table 13
NOMINAL SAMPLING REQUIREMENTS FOR VARIOUS FIELD QUANTITIES

<u>Quantity</u>	<u>Value</u>
N_s , total number of spatial samples (per scalar unknown)	$-(L/\Delta L)^d = (2\pi L/\lambda)^d$
N_t , number of time steps for time-domain model	$-(L/\Delta L) = (2\pi L/\lambda)$
N_f , number of frequency steps to characterize spectral response from frequency-domain model	$-(L/2\Delta L) = N_t/2$
N_{ms} , number of excitations sources for monostatic radar cross section in one plane*	$-(4L/\Delta L) = 8\pi L/\lambda$
N_{bistat} , number of far fields needed for bistatic pattern in one observation plane*	$-N_{ms} = (4L/\Delta L)$

where:

λ is the wavelength at the highest frequency of interest.

ΔL is the spatial resolution being sought.

L is object maximum object dimension or dimension in observation plane.

d is the number of spatial dimensions being sampled, and is not necessarily the problem dimensionality D . The distinction is important because when an appropriate Green's function is available, the source integrals are usually one dimension less than the problem dimension, i.e., $d = D-1$. An exception is an inhomogeneous, penetrable body where $d = D$ when using an integral equation.

*Assuming ~ 6 samples per lobe of the scattering pattern are needed.

Table 14
SOME ANALYTICAL MEANS OF REDUCING COMPUTER TIME

<u>Method</u>	<u>Implementation Requirements</u>	<u>Motivation</u>
Specialized Green's functions	Delta-source fields for geometry of interest	Eliminate need to solve unknowns over the specific surface
Hybrid models	Interactions between parts of problem modeled using different formulations	Model self interaction using most efficient formulation
Impedance boundary conditions	Boundary smooth enough that ratio of local fields determined by constitutive parameters	Reduce number of unknowns by a factor of two
Physical optics	Local surface field determined by constitutive parameters and incident field	Eliminate all unknowns, but excitation dependent

Table 15
A SUMMARY OF SOME SPECIALIZED GREEN'S FUNCTIONS

<u>Problem</u>	<u>Green's Function</u>	<u>Characteristics</u>
Planar, penetrable interface.	Continuous radial spectrum [Banos (1966)].	Infinite integrals whose numerical evaluation is time consuming and having analytical approximations of limited applicability.
Infinite, circular cylinder.	Continuous axial spectrum and discrete azimuthal (cosine/ sine) series [Lucke (1951)].	Adds to difficulty of integral evaluation needed to sum azimuth series.
Sphere.	Discrete spectrum in elevation (Legendre polynomials) and azimuth (cosine/sine series) [Tesché & Neureuther (1970)].	Double series can be computationally demanding. Simplifies for single, radial monopole.
Infinite, parallel plates.	Discrete spectrum (infinite set of images) in transverse direction [Tesché (1972)].	Series is slowly convergent, and is poorly behaved at resonant separations.
Rectangular waveguide.	Extension of parallel plate treatment to two transverse image sets and double series.	Double series increases the convergence problem from $\sim N$ to $\sim N^2$.
Rectangular cavity.	Extension of rectangular waveguide to closed cavity, with three image sets and triple series [Wu and Chang (1986)].	Triple series convergence $\sim N^3$, but transformed to single series by Wu and Chang.

Table 16
SOME HYBRID MODELS USED IN COMPUTATIONAL ELECTROMAGNETICS

<u>Approach</u>	<u>Implementation</u>	<u>Motivation</u>
GTD/IE [Thiele and Newhouse (1969)].	Addition to MoM impedance matrix of fields due to diffracted rays.	Reduce number of MoM unknowns when modeling large objects.
Modal/DE [Mei (1974)].	Develop solution for penetrable object using DE and match to modal outward propagating fields.	To include radiation condition in frequency-domain, DE model.
Modal/Experimental [Wacker (1981)].	Measure antenna near fields over plane, cylinder, or sphere and transform to far field using modal expansion.	To permit measurement in near field of antennas too large for practical, direct far-field measurement.
IE/DE [Taflöv and Umashankar (1980)].	Apply IE model over enclosing surface within which fields are modeled using DE.	To exploit exterior-region advantage of IE and interior-region generality of DE.

Table 17
SOME NUMERICAL MEANS OF REDUCING COMPUTER TIME

Method	Implementation Requirements	Motivation
Iterative technique to solve direct matrix	Various special-purpose iteration routines which include convergence checks	Reduce solution time from N_s^3 to $\sim N_s^2$
Exploiting problem symmetries.	Special programs for filling and solving symmetric systems of various kinds.	Save varying amounts of computer time/storage, but limited in applicability.
Near-neighbor approximations.	Strategy for setting interaction-window width and exploiting sparseness of direct matrix.	Save factor of N_s or more in time/storage, but produces approximate solution.
Adaptive modeling.	Variation of numerical model as computation proceeds to achieve a specified local accuracy.	Fewer unknowns and controllable error, but increased programming complexity.
Model-based parameter estimation.	Development of approximate, simpler expressions to replace rigorous, but computationally time-consuming, rigorous ones.	Reduce time needed to obtain direct matrix, or number of samples needed, e.g., to estimate transfer function.
Numerical Green's function	Development and storing of solution matrix for primary object of interest.	Need only to solve interaction terms to model effect of other nearby objects.

Table 18
ERROR TYPES THAT OCCUR IN COMPUTATIONAL ELECTROMAGNETICS

CATEGORY	DEFINITION
Physical Modeling Error, ϵ_p	Arises because the numerical model used is normally an idealized mathematical representation of the actual physical reality.
Numerical Modeling Error, ϵ_N	Arises because the numerical results obtained are only approximate solutions to that idealized representation, and is comprised of two components: <ol style="list-style-type: none"> 1) Solution error—The difference that can exist between the computed results and an exact solution even were the linear system of equations to be solved exactly, due to using a finite number of unknowns; and 2) Equation error—The equation mismatch that can occur in the numerical solution because of roundoff due to finite-precision computations, or when using an iterative technique because of limited solution convergence.

Table 19
GENERIC GUIDELINES FOR IE MODELING

MODELING PARAMETER OR ISSUE	NOMINAL RANGE OR VALUE	TYPICAL REASON
Wire Length, L	$L > 10d$	Neglect of end caps in thin-wire treatment.
Wire diameter, d	$\lambda > \pi d$	Neglect of circumferential effects.
Wire segment length, Δ As related to diameter	$\Delta > d$	Use of thin-wire kernel in integral equation. Can be relaxed by use of extended kernel [Burke and Poggio (1980)].
As related to wavelength	$\Delta < \lambda/2\pi$	Necessity of sampling current densely enough per unit wavelength.
Step change in wire radius, δa	$\Delta > 10\delta a$	Neglect of sources on stepped surface (similar to end-cap problem).
Source location	Do not place on open-ended segment	Avoids non-physical situation of driving wire at open end.
Angle of wire bend, α	$\alpha > 2\pi a/\Delta$	Keeps adjacent wires from occupying too large a common volume.
Axial separation of parallel wires, r With match points aligned	$r > 10a$	Avoids placing one match point (when using delta-function weights) in error field of another
Otherwise	$r > 3a$	Neglect of circumferential current variation.
Wire mesh model of solid surface Mesh size $\Delta \times \Delta$	$\Delta < \lambda/10$	To reduce field "leakage" to acceptable level.
Wire radius	$a \approx \Delta/2\pi$	To have wire area equal to surface area of solid.
Surface-patch area, Δ_s Frequency domain	$\lambda > 2\pi\sqrt{\Delta_s}$	Need to sample currents densely enough in wavelengths.
Time domain	$\lambda_{min} > 2\pi\sqrt{\Delta_s}$	densely enough in
Piecewise model of curved wire or surface of radius of curvature R	$R > \Delta$ $R > \sqrt{\Delta_s}$	Necessity of sampling a circular arc at least 6 times per 2π radians.
Starting time in time-domain solution, t_s	$E^{inc}(t) _{max} \geq 10^6 E^{inc}(t_s)$	To achieve numerically smooth buildup of exciting field and 10^6 accuracy
Stopping time in time-domain solution, t_p	$E^{inc}(t) _{max} \geq 10^6 E^{inc}(t \rightarrow \infty) - E^{inc}(t_p) $ or $I(t)$ reaches steady state	Stabilize final response.
Time step in TD solution, δ	$c\delta \leq \Delta$	Satisfies Courant stability condition. Required for explicit solution.
Maximum frequency of transient source in TD solution, using Gaussian excitation, i.e. $E^{inc}(t) = \exp(-g^2 t^2)$	$g \sim 2f_{max}$	Ensures source spectrum does not exceed upper frequency for model validity.

Table 20
INTERNAL CHECKS USEFUL AS MEASURES OF SOLUTION VALIDITY

CONVERGENCE MEASURES			
Measure	Example	Tests	Properties
Local	$\lim[I(s)], \lim[E(r)]$ as $N \rightarrow N_{\max}$	Convergence of input impedance, current, fields, etc.	Reasonable measure of solution behavior, but can yield non-monotonic result.
Global	$\int I(s)I'(s)ds$ or $\int E(r)E'(r)dx^*$	Convergence over entire object of current or convergence of field in $n=1, 2$, or 3 dimensions.	A more complete measure of convergence.
Random (local or global)	$\Sigma F(r_s)$, with $F(r_s)$ a field quantity which is function of a random variable r_s .	Convergence of any field quantity measured by a random observation variable.	Permits estimation of convergence and uncertainty of convergence estimates.
OTHER MEASURES			
Power Balance	$P_{in} + P_{loss} = P_{radiated}$	Whether supplied power equals sum of radiated plus dissipated power.	Provides good check on antenna source model for radiation resistance. A necessary, but not sufficient condition.
Boundary Condition Matching	$E_{tan}(r') = 0$, r' on object	Degree to which specified conditions on the boundary modeled are satisfied.	Most fundamental check on solution. Consistency requires use of same weight function as for model itself. Can be computationally expensive. Necessary and sufficient condition.
Reciprocity	$E(j_1, j_2^{scat}) = E(j_2^{inc}, j_1^{scat})$	Whether interchanging observation and source locations yields identical results.	Useful check for antenna and bistatic scattering patterns. Necessary but not sufficient condition.
"Non-physical Behavior" of Solution	—	Whether computed results exhibit physically reasonable behavior.	Can be a subjective check. One example is provided by spatial oscillation in current when thin-wire approximation is violated.

Table 21
EXTERNAL CHECKS USEFUL AS MEASURES OF SOLUTION VALIDITY

Measure	Example	Tests	Properties
ANALYTICAL			
Any observable provided by a formally exact solution.	Sources, near and far fields.	Any observable provided by the computer model.	Provides a necessary and sufficient condition for solution validity. Available for only special geometries, but gives canonical benchmarks.
COMPUTATIONAL			
Far fields.	Radiation pattern, bistatic and monostatic scattering pattern.	Consistency of the quantity least sensitive to solution errors.	A useful test, but one which is often subject to angle shifts between results from two models.
Near fields and sources.	Near-field cuts, current and charge distributions.	Quantities most often directly computed by model.	A more demanding test for comparison, but one which often exhibits spatial shifts between models.
Impedance/susceptance.	—	Source models and single-port input characteristics.	Especially sensitive measure in terms of input susceptance. Highly advisable to examine over a range of frequencies because shifts in frequency also occur.
EXPERIMENTAL			
Same observables as used for computational checks.	—	Physical modeling error and relative correlation of actual problem with numerical model.	Perhaps the most reassuring check to make, but also often the most difficult.

RECENT ADVANCES TO NEC: APPLICATIONS AND VALIDATION *

G. J. Burke
Lawrence Livermore National Laboratory
P.O. Box 5504, L-156, Livermore, CA 94550

ABSTRACT

Capabilities of the antenna modeling code NEC are reviewed and results are presented to illustrate typical applications. Recent developments are discussed that will improve accuracy in modeling electrically small antennas, stepped-radius wires and junctions of tightly coupled wires, and also a new capability for modeling insulated wires in air or earth is described. These advances will be included in a future release of NEC, while for now the results serve to illustrate limitations of the present code. NEC results are compared with independent analytical and numerical solutions and measurements to validate the model for wires near ground and for insulated wires.

1. INTRODUCTION

Computer modeling of antennas, since its start in the late 1960's, has become a powerful and widely used tool for antenna design. Computer codes have been developed based on the Method of Moments, Geometrical Theory of Diffraction, or integration of Maxwell's equations. Of such tools, the Numerical Electromagnetics Code - Method of Moments (NEC) [1] has become one of the more widely used codes for modeling resonant sized antennas. There are several reasons for this, including the systematic updating and extension of its capabilities, extensive user-oriented documentation and accessibility of its developers for user assistance. The result is that there are estimated to be several hundred users of various versions of NEC world wide.

NEC has been under development for more than ten years (in earlier forms it was known as BRACK and AMP.) It is a hybrid code which uses an electric field integral equation to model wire-like objects and a magnetic field integral equation to model closed surfaces with time harmonic excitation. NEC is commonly applied to modeling antennas in VLF to VHF applications on ships, vehicles or on the ground. It includes a number of features for efficient modeling of antennas and scatterers in their environments including antennas interacting with or buried in a finitely conducting ground. Recent work on NEC has added a capability for modeling insulated wires in the air or ground and improved accuracy for electrically small antennas and the treatment of an abrupt change in wire radius.

We will review the current status of NEC and discuss recent enhancements and new developments. Some typical results and validation examples are included.

2. THE BASIC MODELING APPROACH

NEC combines an electric-field integral equation (EFIE) for wires and a magnetic-field integral equation (MFIE) for surfaces to model complex structures. Each equation has advantages for particular structure types. The EFIE is well suited for thin wires with small or vanishing conductor volume, while the MFIE, which fails for thin wires, is more attractive for closed, voluminous structures, especially those having large smooth surfaces. The EFIE can also be used to model surfaces and is preferred for thin structures where there is little separation between a front and back surface. Although the EFIE is specialized to thin wires in NEC, it can be used to represent surfaces as wire grids. Grid models are often more accurate for far-field quantities than for surface fields or currents, however. For a structure containing both wires and surfaces the EFIE and MFIE are coupled, as was described by Albertsen, Hansen and Jensen in [2]. A derivation of the EFIE and MFIE used in NEC is given by Poggio and Miller in [3].

The thin wire approximation is applied to the EFIE to reduce it to a scalar integral equation. The basic approximations are to neglect transverse currents and transverse variation of the axial current on the wire, and to enforce the boundary condition on electric field in the axial direction only. In NEC, the current is lumped as a filament on the wire axis and the boundary condition is enforced on the wire surface to facilitate evaluation of the kernel. These approximations are valid as long as the wire radius is much less than the wavelength and much less than the wire length. An alternate kernel for the EFIE, based on an extended thin-wire approximation in which the current is a tubular distribution on the wire surface, is also included for wires having too large a radius for the thin-wire approximation

* Work performed under the auspices of the U. S. Department of Energy by the Lawrence Livermore National Laboratory under Contract W-7405-Eng-48.

[4]. With the thin-wire approximation, the EFIE for a wire with radius $a(s)$ becomes

$$-\hat{s} \cdot \vec{E}^i(\vec{r}) = \frac{-j\eta}{4\pi k} \int_C I(s') \left(k^2 \hat{s} \cdot \hat{s}' - \frac{\partial^2}{\partial s \partial s'} \right) g(\vec{r}, \vec{r}') ds', \quad \vec{r} \in C(s) \quad (1)$$

where

$$g(\vec{r}, \vec{r}') = \exp(-jkR)/R, \quad R = [|\vec{r} - \vec{r}'|^2 + a^2(s)]^{1/2}$$

$$k = \omega(\mu_0\epsilon_0)^{1/2}, \quad \eta = (\mu_0/\epsilon_0)^{1/2}.$$

I is the induced current, \vec{E}^i is the exciting field, \hat{s} and \hat{s}' are unit vectors tangent to the wire at s and s' , and \vec{r} and \vec{r}' are vectors to the points s and s' on the wire contour C .

The MFIE for a closed surface S is

$$-\hat{n}(\vec{r}) \times \vec{H}^i(\vec{r}) = -\frac{1}{2} \vec{J}_s(\vec{r}) + \frac{1}{4\pi} \int_S \hat{n}(\vec{r}') \times [\vec{J}_s(\vec{r}') \times \nabla' g(\vec{r}, \vec{r}')] dA', \quad \vec{r} \in S \quad (2)$$

where $\hat{n}(\vec{r})$ is the unit vector normal to the surface at \vec{r} , \vec{H}^i is the exciting magnetic field and \vec{J}_s is the surface current. Eq. (2) is derived from the condition that magnetic field is zero inside a perfectly conducting solid, with a jump of $\vec{J}_s \times \hat{n}$ at the surface. Hence it is valid only for a closed perfectly conducting shell. Eq. (2) is separated into coupled scalar integral equations for components of the surface current. For models involving both wires and surfaces the electric field due to surface currents is included in the right-hand side of Eq. (1) and the magnetic field due to wires is included in the right-hand side of Eq. (2) leading to coupled integral equations for the currents.

While Eqs. (1) and (2) apply to an antenna in free space, they are easily extended to an infinite dielectric or conducting medium by replacing ϵ_0 and μ_0 by ϵ and μ of the medium. When the antenna is located near a ground plane, the right-hand sides of Eqs. (1) and (2) are modified to include the increment in the field of the source current due to the interface, either using an image or (for wires only in NEC-3) the Sommerfeld-integral solution for a point source near the interface.

The integral equations (1) and (2) are solved numerically by a form of the method of moments, which involves expanding the unknown current in a summation of basis functions and enforcing equality of weighted integrals of the fields to reduce the integral equation to a matrix equation [5]. The weighting functions for both wires and surfaces are chosen to be delta functions, resulting in a point sampling of the fields known as the collocation method. Wires are divided into short straight segments with a sample point at the center of each segment, while surfaces are approximated by a set of flat patches or facets with a sample point at the center of each patch.

Delta functions are also used as the current expansion functions on surfaces. For the MFIE this elementary Galerkin's method has been found to provide good accuracy on large smooth surfaces. Due to the nature of the integral-equation kernels, however, the choice of current expansion functions is more critical in the EFIE than in the MFIE.

The expansion functions for the current on wires are chosen so that the total current on segment number j has the form

$$I_j(s) = A_j + B_j \sin k_s(s - s_j) + C_j \cos k_s(s - s_j), \quad |s - s_j| < \Delta_j/2 \quad (3)$$

where s_j is the value of s at the center of segment j and Δ_j is the length of segment j . The factor k_s is normally equal to the wave number in the medium containing the wire. This expansion was first used by Yeh and Mei [6] and has been shown to provide rapid solution convergence. In NEC, the constants in Eq. (3) are chosen so that the basis functions are generalized B-splines extending over three segments (or more at multiple wire junctions) and constructed from the three-term sinusoidal function.

Two of the three constants in Eq. (3) are eliminated by continuity conditions on current and charge at the segment ends. Current is continuous or satisfies Kirchhoff's current law at multiple wire junctions. Charge is continuous on uniform wires. At a junction where the wire radius changes, the linear charge density (or $\partial I/\partial s$) on the wire with radius a_i is made proportional to

$$\Psi_i = \left[\ln \left(\frac{2}{ka_i} \right) - 0.5772 \right]^{-1}. \quad (4)$$

This condition on charge was derived in [7] by analysis of a wire with tapered radius, and hence is often called the Wu-King condition. On a wire passing from air into ground, the charge is discontinuous by the ratio of media permittivities [8]. Expressions for the basis functions that result from applying these conditions on current and charge are included in [1].

Where a wire connects to a surface, a more realistic representation of the surface current is needed than the delta-function expansion normally used with the MFIE. The treatment used in NEC is similar to that used by Albertson et al. [2]. In the region of the wire connection, the surface current includes a singular component with the form $\vec{\rho}/\rho^2$, where $\vec{\rho}$ is the surface vector from the attachment point, as well as a continuous function providing interpolation to the current on adjacent patches.

As a result of the continuity conditions on current and charge there remains one unknown associated with each basis function to be determined by solving the electromagnetic interaction equations. Enforcing appropriate conditions on the current and charge results in rapid convergence of the solution. The danger is that the conditions may be inappropriate under some conditions, as may occur in NEC at a change in wire radius and at junctions of closely spaced wires. The result is then slower convergence. More accurate conditions for these cases are being developed, as discussed in Section 4.

3. CAPABILITIES OF NEC

Two versions of NEC are currently being released: NEC-2 which has no distribution restrictions and NEC-3 which is classified as Defense Critical Technology, and hence can be released only to DoD agencies and contractors. Foreign release of NEC-3 requires specific DoD approval. The only difference in the capabilities of these codes is that NEC-3 can model wires that are buried or penetrate from air into the ground, while NEC-2 is limited to antennas in free space or above a ground plane. The capabilities of these codes are summarized below.

3.1 Source Modeling

A voltage source on a wire may be modeled by an applied field on a segment or a discontinuity in charge between segments. The charge discontinuity, or "bicone" source model yields a more localized excitation than the field over a segment, but it is accurate only for small wire radius [9]. Alternatively, a structure may be excited by a plane wave with linear or elliptic polarization or by the near field of an infinitesimal current element.

3.2 Nonradiating Networks, Transmission Lines and Loading

Nonradiating two port networks and transmission lines may connect points on wires. These are modeled by constructing a driving-point admittance matrix from the admittance matrix of the entire structure to avoid modifying the larger matrix. Lumped or distributed RLC loads may be specified on wires. Also, the conductivity of a round wire may be specified and the impedance computed taking account of skin depth.

3.3 Ground Effects

Three options are available for an antenna in or near the ground. A perfectly conducting ground is modeled by including the image field in the kernel of the integral equation. This doubles the time to compute the interaction matrix. An approximate model for an antenna over a finitely conducting ground uses the image modified by Fresnel plane-wave reflection coefficients. This approximation is usable for antennas at least 0.1 to 0.2 wavelengths above the ground and doubles the time to fill the matrix.

NEC-3 includes an accurate treatment for wire structures above, below, or penetrating the ground surface. The solution is based on the Sommerfeld-integral formulation for the field near the interface. For the moment-method solution, the field values are obtained by table lookup and parameter estimation involving a model for the functional behavior of the field transmitted across the interface. The current expansion is modified to account for the discontinuity in charge on a wire penetrating the interface. This numerical treatment for buried wires is described in [8]. NEC-2 includes a similar ground model limited to wires above the ground.

Over ground, NEC will compute the radiated space wave, with direct and reflected rays, and can also include the surface wave. A cliff or abrupt change in ground parameters is treated by computing the reflected ray for the appropriate ground parameters and height at the reflection point, but without including diffraction. Hence at grazing angles to the ground, since the reflection point is at an infinite distance, only the outer medium affects the field. The inner medium is always used in computing the antenna currents. Also, no model for a cliff is available when the surface wave is computed. Options for a more complete treatment of a cliff are discussed in Section 4.5.

3.4 Efficient Solution Methods

The matrix equation that results from the moment method is solved by LU decomposition, with the factors saved for reuse when the excitation or other parameters that do not alter the matrix are

changed. Rotational symmetry or reflection symmetry in one to three planes can be used to reduce the times to fill and to factor the matrix. New wires or surfaces may be added to a structure for which the matrix has already been computed, factored and saved on a file. The new solution is found from the self and mutual interaction matrices through a partitioned-matrix algorithm with no unnecessary repetition of calculations for the basic structure. This feature can be used to take advantage of symmetry in a portion of a structure to which unsymmetric parts connect or interact.

3.5 Input

A user-oriented input scheme permits defining straight wires, arcs, and surfaces. Shifts, rotations, and reflections can be used in building complex structures. Electrical connections are determined in the program by searching for wire ends and patch centers that coincide. A solution can be repeated for modified model parameters (transmission lines, loading, ground, etc.) in the input sequence without respecifying parameters that are not changed.

3.6 Output

Output may include current, charge density on wires, input impedance, admittance, input power, ohmic loss and efficiency. For transmitting antennas, radiated field and power, power gain, directive gain, average gain and ground-wave field are available. Near \vec{E} and \vec{H} can be computed at points on or off of the antenna. Receiving patterns and scattering cross sections are available with incident field excitation.

3.7 Modeling Guidelines

There are a number of rules that must be followed in specifying the model in order to obtain accurate results from a code such as NEC. Unfortunately many of these rules can be ignored and the code will still give results — usually wrong. Hence a user should become familiar with the modeling guidelines before attempting to use the code. Some of the guidelines are summarized below.

For accurate results, the lengths of wire segments should be less than about 0.1λ . Longer segments up to about 0.14λ may be acceptable on long, straight wires or noncritical parts of a structure, while shorter segments, 0.05λ or less, may be needed in modeling critical regions of an antenna. There is also a minimum segment length set by the computational precision. With seven place precision, common on 32-bit computers, the minimum segment length is about $10^{-4} \lambda$. In double precision the minimum length is about $10^{-8} \lambda$. The restrictions on minimum segment length will be relaxed in a future version of the code, as discussed in Section 4.1.

The wire radius a relative to λ is limited by approximations used in the kernel of the EFIE. Two options are available in NEC: the thin-wire kernel (TWK) and the extended thin-wire kernel (ETWK). With the ETWK the ratio of segment length to radius can be as small as 0.5 before instabilities appear in the solution, while with the TWK the limit is 2. In either case, only currents in the axial direction on a segment are considered, and there is no allowance for variation of the current around the wire circumference. Also, the ETWK, when it is selected, works only on straight sections of wire. The acceptability of these approximations depends on both the value of a/λ and the tendency of the excitation to produce circumferential currents or current variation. The capacitive effect of a step in radius, and consequent shortening of the electrical length, is not modeled accurately in the present NEC. If a discontinuous radius must be modeled it is best to keep the ratio of radii less than two, and to use segments as long as possible at the discontinuity. A new model for a stepped-radius wire is described in Section 4.3.

Electrically small loops present special accuracy problems. In single precision the minimum loop circumference is about 0.07λ while in double precision it is about $3(10^{-4}) \lambda$. These limits can be removed, as described in Section 4.2, by including loop basis and weighting functions on each small loop. This capability is not yet in the released version of NEC, however. For a loop over a finitely conducting ground the solution fails when the loop size is somewhat less than resonant. Hence small loops over ground cannot presently be modeled.

Due to the nature of the MFIE, surfaces modeled with patches must be closed. A reasonable maximum for the area of a surface patch on a smooth surface appears to be 0.04 square wavelengths.

An important consideration in using NEC is the solution time versus model size since this may limit the amount of detail that can be modeled and the segment and patch densities. For a model using N wire segments, the solution time in seconds on a CDC 7600 computer is approximately

$$T = 3(10^{-4})kN^2/M + 2(10^{-6})N^3/M^2 \quad (5)$$

where M is the number of degrees of symmetry and k is 1 for free space or any infinite medium, 2 for an antenna over perfectly conducting ground or real ground modeled with the reflection coefficient approximation, and 4 to 8 for the accurate (Sommerfeld integral) treatment of finitely conducting ground. The first term in Eq. (5) is the time to fill the interaction matrix and the second term is the time for the LU decomposition.

For a model using N_p patches, the solution time is about

$$T = (10^{-5})k(2N_p)^2/M + 2(10^{-6})(2N_p)^3/M^2$$

since two rows and columns in the matrix are associated with each patch. The reduced time factor for filling the matrix, due to the delta-function current expansion, is one attractive feature of this solution method. A more complete relation for running time is contained in [1], Part III.

When the Sommerfeld ground treatment is used, a fixed time of about 15 seconds on a CDC 7600 computer is needed to generate the interpolation tables. The tables depend only on the ground parameters and frequency, however and can be saved for reuse in any case in which these parameters are the same.

4. NEW DEVELOPMENTS IN NEC

Although the present NEC offers a reasonably versatile and accurate modeling capability, there are a number of situations that may not be handled accurately. Recent work on the thin-wire code has increased the precision and model accuracy for electrically small antennas and developed more accurate treatments for wires with discontinuous radius and junctions of tightly coupled wires. Also, a model for insulated wires has been added to a version of NEC-3. These developments, which will be included in future versions of NEC, are described below.

4.1 Enhanced Precision for Electrically Small Structures

NEC was originally developed on computers having 60-bit numerical precision, so little attention was given to preserving numerical accuracy. Problems may become apparent when the code is run on computers with 32-bit single precision. Precision loss is particularly severe when modeling electrically small antennas, and often forces the use of double precision. Hence we have NEC-2S and NEC-3S codes in single precision and NEC-2D and NEC-3D in double precision on DEC/VAX computers. Double precision, of course, increases the storage requirement and more than doubles the solution time, so some large VLF models are run in single precision because double is unaffordable. Problems in modeling electrically small antennas can be divided into those that affect any small structure and those that occur with electrically small loops. Changes to correct low frequency problems other than loops are described in [10] and are summarized below.

To improve the accuracy of the single precision NEC, we have revised the code for evaluation of the basis functions and fields. A basic problem at low frequencies is the cancellation between the constant and $\cos(k_s s)$ components in Eq. (3). In the new code, the same spline basis functions are used but they are evaluated in the form

$$I_j(s) = A'_j + B_j \sin k_s(s - s_j) + C_j[\cos k_s(s - s_j) - 1], \quad |s - s_j| < \Delta_j/2$$

over each segment. To permit modeling extremely short segments, it was also necessary to arrange the order of calculation to avoid overflow or underflow of intermediate results in evaluating the spline functions.

Precision loss is also a problem in the field evaluation. The present code evaluates the electric field of each segment, including the contributions of point charges at the segment ends that result from the current dropping abruptly to zero. These point-charge terms dominate the field of a segment at low frequencies but are cancelled by adjacent segments when current is continuous. Since continuity of current is ensured in the current expansion, these terms can be dropped from the field evaluation to preserve accuracy.

The fields due to $\sin(k_s s)$ and $\cos(k_s s)$ currents are evaluated from exact closed-form expressions when k_s is equal to the wave number k in the surrounding medium, while the field of a constant current involves an integral that must be evaluated numerically or by approximation. Even with the point-charge terms removed, the field equations suffer severe cancellation that necessitates the use of series expansions for small kR , R much greater than segment length and for ρ much less than z , where ρ and z are the coordinates of the evaluation point relative to a segment on the z axis and $R = (\rho^2 + z^2)^{1/2}$.

To compute the input resistance of small antennas it was necessary to maintain the accuracy of the real part of the electric field although, for some components, it decreases faster than the imaginary part by a factor of $(kR)^3$.

These changes have been made in a code called NEC3VLF. This is an interim code to which additional changes are being made. However, a comparison of results from NEC-3S, NEC-3D and NEC3VLF reveals some of the limitations of NEC-3 and the potential improvements. One way to determine the low frequency limit of the code is to decrease the frequency until the solution obviously fails. Such results are shown for a dipole in Fig. 1. NEC-3S fails at a point that could cause problems in modeling practical VLF antennas. NEC-3D could handle most practical applications, but at the expense of increased solution time and storage. NEC3VLF is limited only by the underflow of the conductance and is also faster than NEC-3S due to a faster series approximation for the integral of $\exp(-jkR)/R$.

The solution in NEC-3S can also become inaccurate for a fixed length dipole as the number of segments is increased. This problem has been corrected in NEC3VLF as shown in Fig. 2.

4.2 Electrically Small Loops

The difficulty in modeling loops can be seen when the electric field of a wire is written as

$$\vec{E}(\vec{r}) = \frac{-j}{4\pi\omega\epsilon} \left[\nabla \int_L \frac{e^{-jkR}}{R} \frac{\partial I(s')}{\partial s'} ds' + k^2 \int_L \frac{e^{-jkR}}{R} \vec{s}' I(s') ds' \right]. \quad (6)$$

As frequency is reduced the field of a spline basis function is dominated by the first term in Eq. (6), while the second term decreases as ω^2 . However, the sum of equal basis functions around a loop is a constant current with zero derivative, so that the first term in (6) vanishes. Thus the sum of matrix columns representing the spline basis functions around the loop is much less than the individual columns, indicating an ill-conditioned matrix. With unlimited precision the basis functions used in NEC could handle arbitrarily small loops. In single precision, however, accuracy is quickly lost as frequency is reduced.

The problem is still worse when a loop is excited by coupling to a source such as a dipole with large concentrations of charge. The loop current is then determined by the line integral around the loop of the field of the dipole. At low frequency the dipole field is dominated by the gradient term in Eq. (6) which must vanish in the line integral around the closed loop. Hence the sum of matrix rows representing segments of the loop is much less than the individual rows. This problem in NEC is worse than the degeneracy of the basis function fields since the integral of \vec{E} around the loop is sampled in a relatively crude form as the sum of the fields at the centers of segments. Hence, use of double precision does not help the solution for a loop coupled to charge, and large incorrect loop currents can result. This problem may occur whenever a wire end with high charge density is near a small loop.

The problems in modeling small loops can be fixed, as shown in [11], by changing the current expansion to include a function that is constant around the loop, and also using one constant weighting function. This use of loop basis and weighting functions was suggested by D. R. Wilton [12] for a Galerkin moment-method code similar to MININEC [13]. On a small loop, the loop basis function represents the dominant part of the solution while the remaining spline functions account for the small variations of the current. Loop basis functions can also be used on electrically large loops since replacing one spline basis function with a constant loop function does not change the space spanned by the basis. The constant function is exactly equal to an equal-amplitude sum of the spline functions.

In implementing the loop weighting function, the gradient of the scalar potential is dropped from the field evaluation since it must vanish when integrated around a closed loop. With an accuracy and effort compatible with the point matching normally used in NEC, the line integral of \vec{E} around the loop is then approximated from the sum of the vector potentials at the centers of segments. The loop-weighted equation replaces one of the point-matched equations, typically for the same segment on which the spline basis function was removed for the loop basis function. This replacement can introduce some asymmetry on an otherwise symmetric loop, but such effects are within the bounds of convergence of the solution.

A tricky problem in implementing loop basis and weighting functions is that of locating all loops in a complex wire structure. Code has been developed for this task, but it cannot yet handle completely general wire structures.

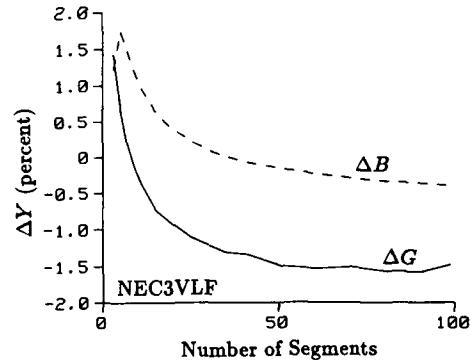
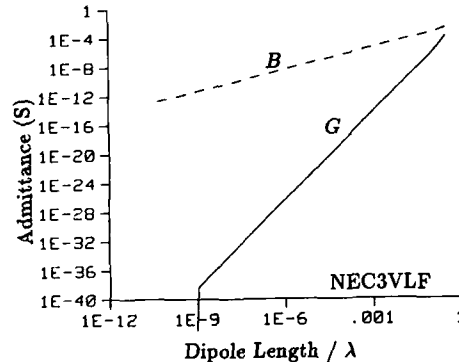
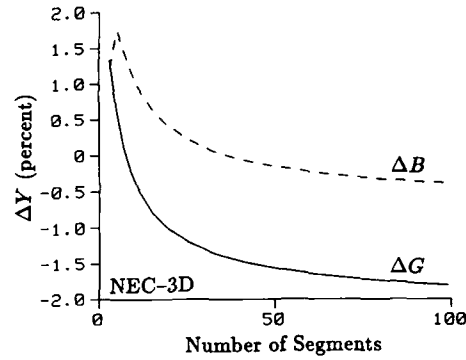
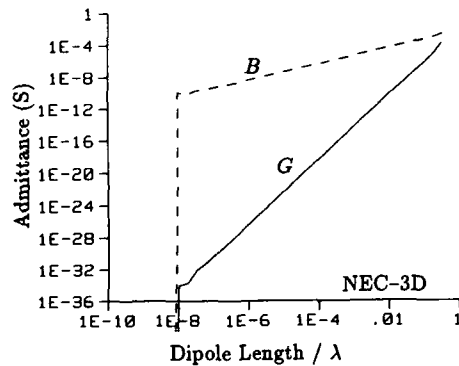
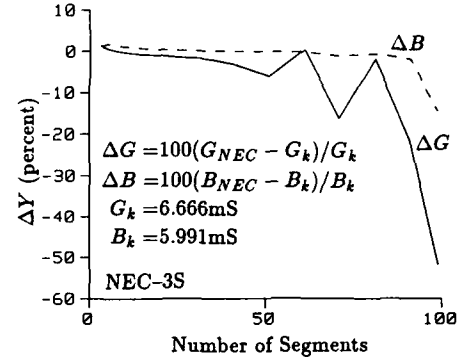
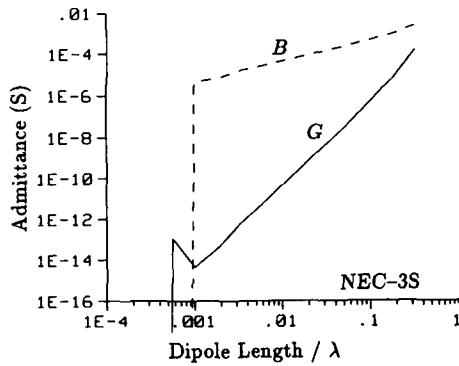


Fig. 1. Input admittance of a dipole antenna computed by NEC-3S, NEC-3D and NEC3VLF. The dipole was modeled with 9 segments and the ratio of radius to dipole length was 10^{-3} . Solution failure is shown by the deviation from the low frequency asymptotic behavior.

Fig. 2. Input admittance of a $\lambda/2$ dipole antenna computed by NEC-3S, NEC-3D and NEC3VLF with varying number of segments. The difference from the King-Middleton result ($G_k + jB_k$) is plotted. The ratio of dipole radius to length is $4.5(10^{-5})$ ($\Omega=20$).

Validation of results for small loops is more difficult than for open vires since the double precision NEC-3, which was used as a standard for dipoles, may give the same wrong results as single precision for loops. The most useful checks on the solution for loops were found to be convergence as the number of segments is increased, the average gain as a check of radiated power versus input power and correction of obviously wrong results for decreasing frequency.

Results for a wire loop containing a voltage source are shown in Fig. 3. In this case the VLF limitation is due to the degeneracy of the fields of the spline basis functions rather than the field sampling. Use of double precision in NEC-3 does reduce the low frequency limit in this case at the expense of

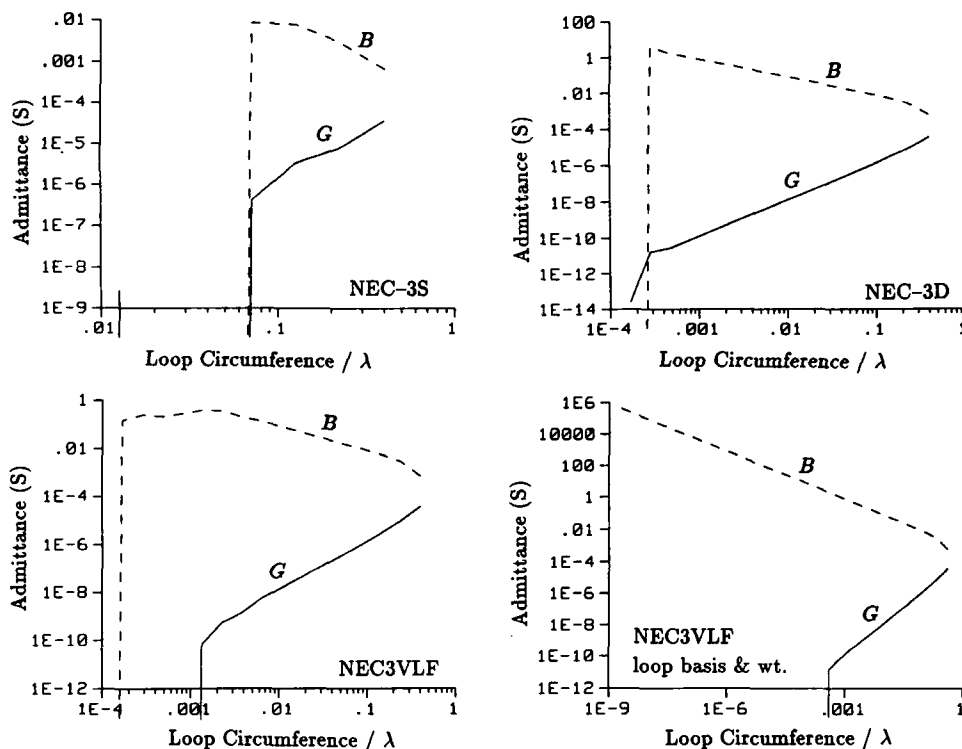


Fig. 3. Input admittance of a loop antenna computed by NEC-3S, NEC-3D and NEC3VLF with and without loop basis and weighting functions. The loop was modeled with 22 segments and the ratio of wire radius to loop radius was 0.042. Solution failure is shown by deviation from the low frequency asymptotic behavior.

increased computation time and storage. NEC3VLF with loop basis and weighting functions gives the correct behavior for input conductance to about the same limiting frequency as double precision NEC-3. The new code appears to have no limitation for computation of input susceptance. Hence after the solution for conductance fails the correct value can be determined by integrating the far-field power on an otherwise lossless antenna. The reason for failure of the conductance at a circumference of about $4(10^{-4})\lambda$ has not been isolated, but may be difficult to avoid given the difference in magnitude from the susceptance.

An example of loops coupled to a charge concentration is provided by the stub antenna on a wire-grid fin, shown in Fig. 4. The NEC-3D result has an incorrect clockwise circulation of current on the grid while NEC3VLF with loop basis and weighting functions produced a uniform flow of current toward the stub. As frequency is reduced the incorrect circulating current from NEC-3 grows as f^{-2} relative to the stub current while the NEC3VLF result remains stable. Results for input impedance and average gain of this structure with decreasing frequency are shown in Table 1. The last two frequencies resulted in division by zero in NEC-3S. The difference of the average gains from the correct value of 1.0 was found to be due to extraneous input power at small, but not negligible, voltages across the segments on which match points were dropped. These errors are within acceptable bounds for most applications, however. The minimum frequency at which NEC-3 yields acceptable accuracy will decrease as the wire end, with its charge concentration, is moved further from the loop.

The severity of the errors with the standard NEC solution for small loops would seem to make the use of loop basis and weighting functions essential. However, the difficulties of locating small loops and of implementing this treatment for arbitrary structures including a ground plane may delay its use.

4.3 Modeling Stepped-Radius Wires and Junctions

The problems with the present NEC model for a stepped-radius wire can be seen by comparing the solution for charge with that from an accurate solution of the surface model. Glisson and Wilton have

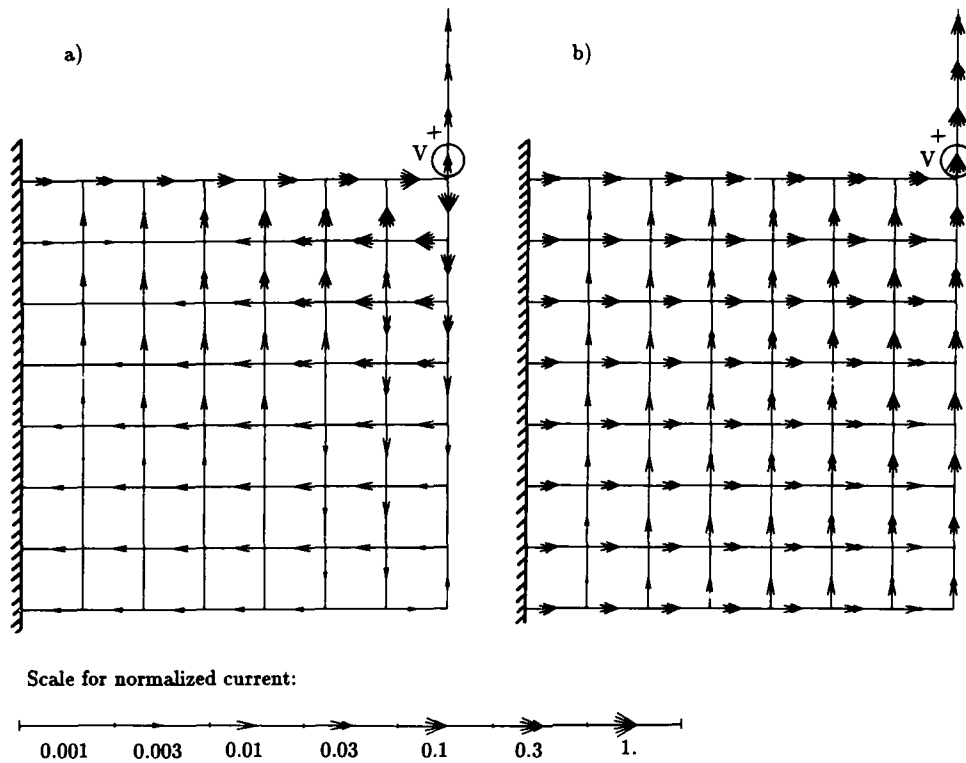


Fig. 4. Imaginary part of current on a wire grid fin with stub antenna. The fin is 0.014λ high and connects to a perfectly conducting ground plane. The source is one volt at the base of the stub and current is normalized by I_{max} ; a) NEC-3D result with incorrect grid current, $I_{max} = 2.62$ mA; b) NEC3VLF with loop basis and weighting functions, $I_{max} = 0.10$ mA.

Table 1. Input impedance and average gain of the wire grid with stub antenna from NEC-3 in single precision (NEC-3S) and double precision (NEC-3D) and NEC3VLF with loop basis and weighting functions. Δ is the width of the openings in the mesh.

Δ/λ	NEC-3S			NEC-3D			NEC3VLF		
	R (ohms)	X (ohms)	\bar{G}	R (ohms)	X (ohms)	\bar{G}	R (ohms)	X (ohms)	\bar{G}
$2 \cdot 10^{-2}$	60.98	$-7.82(10^2)$	1.01	60.98	$-7.82(10^2)$	1.14	$5.91(10^1)$	$-8.00(10^2)$	0.93
$2 \cdot 10^{-3}$	0.60	$-9.38(10^3)$	0.30	$1.78(10^{-1})$	$-9.66(10^3)$	1.19	$1.65(10^{-1})$	$-9.80(10^3)$	0.92
$2 \cdot 10^{-4}$	-7.55	$-8.74(10^4)$	-10^{-4}	$1.76(10^{-3})$	$-9.67(10^4)$	5.66	$1.63(10^{-3})$	$-9.81(10^4)$	0.91
$2 \cdot 10^{-5}$	***	***	***	$1.75(10^{-5})$	$-9.67(10^5)$	484.	$1.63(10^{-5})$	$-9.81(10^5)$	0.91
$2 \cdot 10^{-6}$	***	***	***	$9.81(10^{-5})$	$-9.04(10^6)$	0.08	$1.63(10^{-7})$	$-9.81(10^6)$	0.91

shown in [14], by solution of a surface integral equation for the body of revolution, that the charge density at a step in radius is singular at the outer edge and goes to zero at the inner edge of the step, similar to the behavior of charge on an infinite wedge. If the sinusoidally varying charge density further from the step is extrapolated to the junction, their solution shows the discontinuity given by Eq. (4).

Similar, although probably less accurate, results can be obtained from NEC by modeling the stepped-radius wire as a cage of thin wires. Such a model is shown in Fig. 5 where twelve wires were used to model a monopole with the radius reduced by a factor of two. The monopole in Fig. 5 is viewed at one degree from end on, but is actually very thin relative to its length. As shown by A. Ludwig in [15], the error in modeling a thick wire as a cage of n thin wires is least when the radius of the thin wires is approximately equal to the radius of the thick wire divided by n . Since we did not

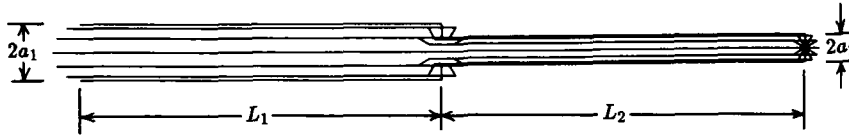


Fig. 5. A twelve-wire cage model of a monopole with stepped radius. The monopole is fed against a ground plane at its thick end. Dimensions are $a_1 = 0.00025\lambda$, $a_2 = 0.000125\lambda$ and $L_1 = L_2 = 0.125\lambda$.

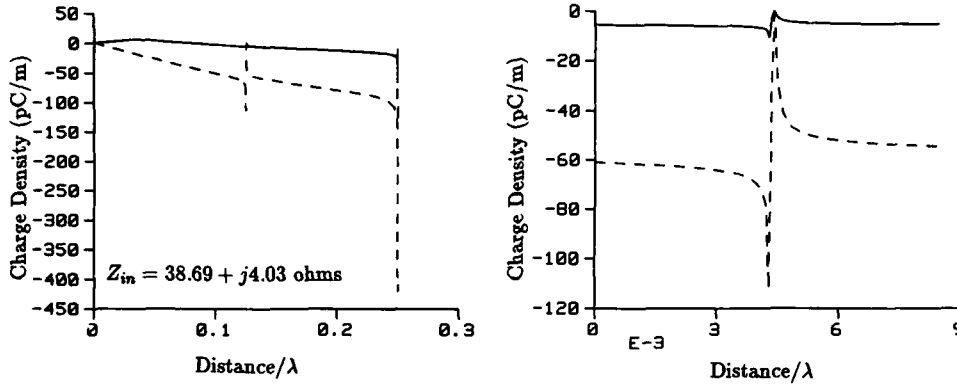


Fig. 6. Charge distribution determined by NEC for the twelve-wire model of a stepped-radius monopole; a) total charge density on the monopole, b) detail of charge density at the step in radius.

want to change the radius of the individual wires, an average radius of $(a_1 + a_2)/24$ was used. In order to resolve the behavior of charge at the step in radius and at the wire end, the short radial wires at these points were modeled with two segments each. The axial segments were tapered to have lengths equal to those of the radial segments at their connections. Results in Fig. 6 from the twelve-wire model show the behaviour of charge at the step and the wire end, and a displacement approximating Eq. (4) at larger distances from the step.

The NEC solution for the charge on the stepped-radius monopole modeled as a single wire is shown in Fig. 7. With segment lengths of $\Delta = 0.03125\lambda$ the discontinuity of Eq. (4) is evident in the charge, and the input impedance is in reasonable agreement with that obtained from the cage model in Fig. 6. However, when the segment lengths at the junction are reduced to $\Delta = 0.0005\lambda$ the charge becomes nearly continuous approaching the step. The discontinuity in charge that is enforced in the current expansion is still present, but it is evident only on the segments immediately adjacent to the step, where it is inappropriate due to the edge effects.

The convergence to continuous charge indicates a failure in the thin wire approximation as implemented in NEC. In NEC-3 and prior codes the current was treated as a filament on the wire axis with match points on the surface to avoid discontinuities in the current, and consequent point charges, at bends and steps in radius. Once the point charges have been dropped from the field evaluation, as discussed in Section 4.1, the current can be located on the wire surface without too much concern for maintaining physical continuity of the distribution. When current is located on the surface and match points on the axis, the solution for charge converges to approximately the discontinuity of Eq. (4) at points away from the edge effects.

Wire ends can then be closed by caps so that the boundary condition matched on the wire axis can represent an exact solution of the surface problem, in the sense of the extended boundary condition [16]. We have used a simple model for end caps, using flat caps with constant charge density that maintains continuity of current and charge with that on the wire. Hence, no new unknowns are introduced by the end caps. Caps were also included on voltage sources to prevent the source field from exciting the inside of the wire. The charge on the source caps is determined by the capacitance of the gap and the source voltage. Hence the field of this charge is added to the excitation vector on the right-hand side of the matrix equation.

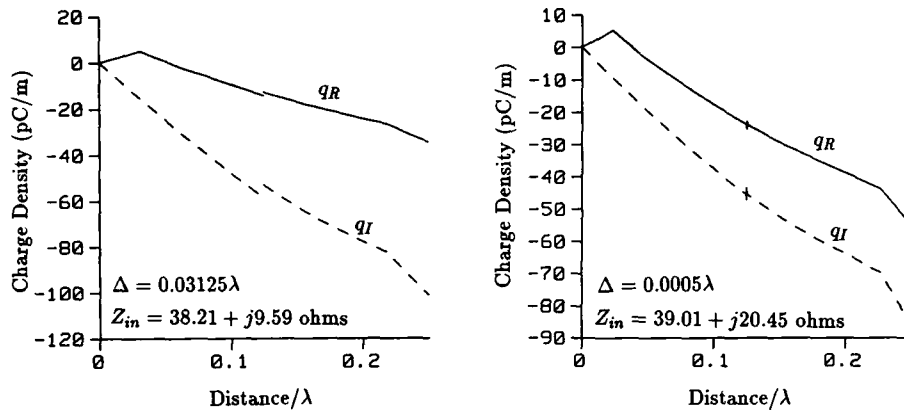


Fig. 7. NEC-3D solution for linear charge density on the monopole of Fig. 5 modeled as a single wire with segment length Δ at the step in radius. For $\Delta = 0.0005\lambda$ segment lengths were tapered for shorter segments at the junction.

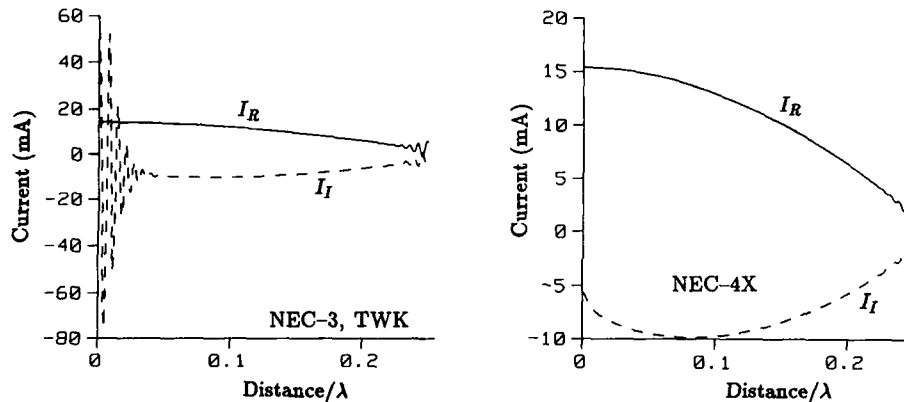


Fig. 8. Current on a $\lambda/4$ monopole excited by a 1 V source at its base, from NEC-3D and NEC-4X. The monopole had a radius of 0.01λ and was modeled with 80 segments.

A condition on charge density that takes into account the proximity of a step in radius and interaction of wires at a junction is determined with minimal computational overhead by executing a small moment-method solution. Any junction on which the charge cannot be determined as uniform due to symmetry is considered isolated from the rest of the structure with the wires extended to infinity away from the junction. An integral equation based on continuity of scalar potential is then solved to determine the distribution of charge. The charge is represented in a piece-wise linear expansion, with triangular basis functions and one semi-infinite basis function on each wire. A single half-triangle and the semi-infinite function were found to be sufficient for convergence.

The new code with the current on the wire surface, end caps, and the new condition on charge at a junction is, for now, called NEC-4X. The effect of including end caps on wire ends and voltage sources is shown in Fig. 8 where results of NEC-3 are compared with NEC-4X. The caps have a significant effect only when the segment length is on the order of the wire diameter or less. However, in view of their effectiveness for short segments, the ETWK will probably be dropped from the next version of NEC. We also plan to investigate the importance of closing the annular surface at a step in radius.

Results of NEC-3 and NEC-4X are compared with the surface solution of Glisson and Wilton for a stepped-radius wire in Fig. 9. The NEC models used relatively short segments at the step. Hence the NEC-3D result shows a large error in current and the nearly continuous charge density seen before. As

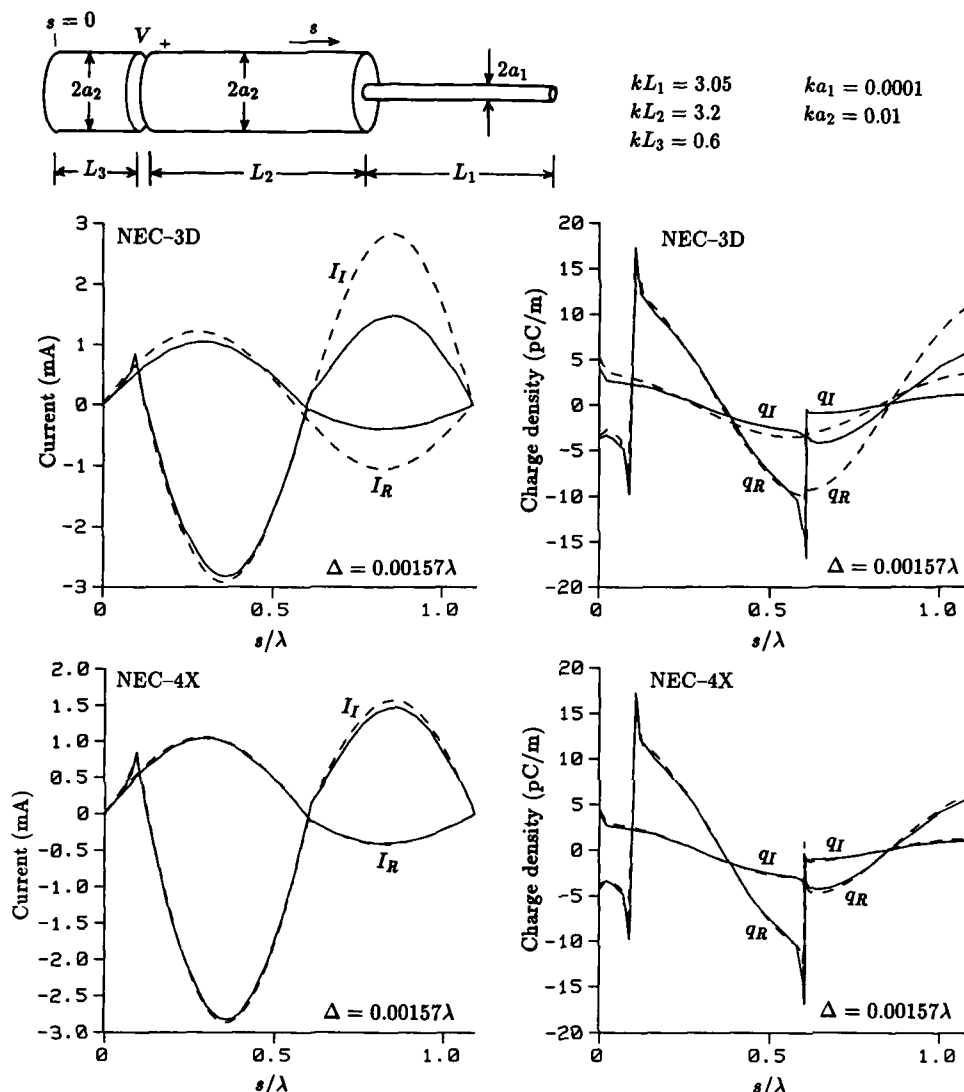


Fig. 9. Current and charge density on a stepped-radius wire. Results from Glisson and Wilton [14] (—) are compared with NEC-3D and NEC-4X (---).

the segment length at the step is increased the error in the NEC-3D result slowly decreases, while the NEC-4X result remains stable and in good agreement with the surface model.

An example of a junction of tightly coupled wires is found in a strip-line EMP simulator constructed from 18 wires as shown in Fig. 10. Results from NEC-3D and NEC-4X for the currents in the wires are shown in Fig. 11 when the line is driven by a one volt source at 4 MHz. NEC-3D is seen to produce nearly uniform currents while results of NEC-4X show larger currents, by nearly a factor of three, in the outer wires as reported from measurements [17]. The vertical component of electric field is compared with measurements in Fig. 12. The field from NEC-3D is too large by a factor of about two, while NEC-4X results are in better agreement with measurements. The results for E_z were numerically integrated from the ground plane to the center of the wire nearest the center plane at $x = 0$ m, $y = 0.02$ m. For a one volt source this integral resulted in 1.93 volts for NEC-3D and 1.05 volts for NEC-4X, further confirming the accuracy of NEC-4X.

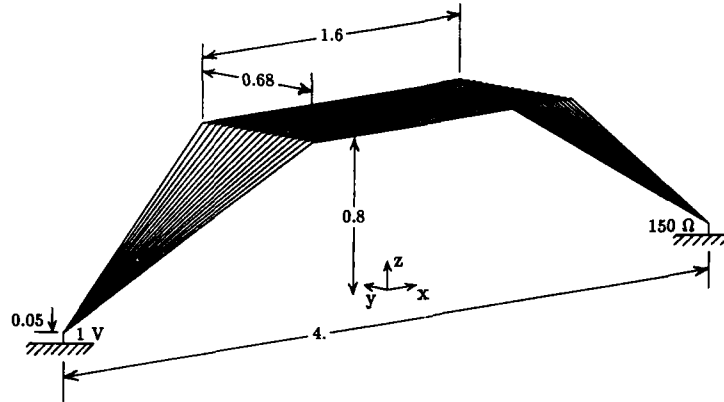


Fig. 10. Strip-line EMP simulator constructed from 18 wires. The line is driven by a voltage source against the ground plane with the other end terminated in a 150 ohm load. Dimensions are in meters, and the wire radius is 0.0008 m.

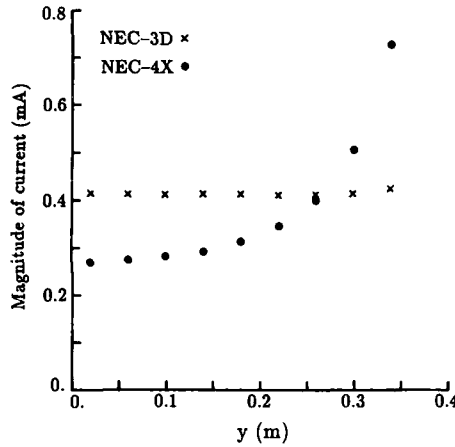


Fig. 11. Magnitude of current on the wires of the strip-line EMP simulator. Current is computed at a point 0.4 m from the center of the simulator toward the source.

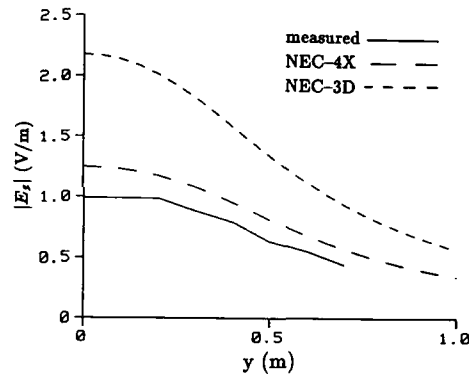


Fig. 12. Vertical electric field in the $z = 0.4$ m plane of the strip-line EMP simulator. NEC-3D, NEC-4X and measured values from [17] are compared.

4.4 A Model for Insulated Wires

A capability for modeling wires with thin insulating sheaths, in air or buried in the earth, has been developed in a version of NEC-3 known as NEC-3I [18]. The field due to radial polarization currents in the sheath is included in enforcing the boundary condition on the wire, as was done by Richmond and Newman in [19]. For a wire with radius a and sheath radius b , as shown in Fig. 13, the insulating sheath with complex relative permittivity $\tilde{\epsilon}_2 = \epsilon_2 - j\sigma_2/\omega\epsilon_0$ in a medium with $\tilde{\epsilon}_1 = \epsilon_1 - j\sigma_1/\omega\epsilon_0$, is replaced by an equivalent polarization current of

$$\vec{J}_s(\rho, z, \phi) = j\omega\epsilon_0(\tilde{\epsilon}_2 - \tilde{\epsilon}_1)[\vec{E}^S(\rho, z, \phi) + \vec{E}^I(\rho, z, \phi)], \quad a \leq \rho \leq b$$

radiating in medium 1. \vec{E}^S is the electric field due both to currents on the wire and to \vec{J}_s itself, and \vec{E}^I is the excitation field. To retain a one dimensional integral equation for the axial wire current, the sheath is assumed to be electrically thin, and the total field in the sheath is assumed to be dominated by the radial field due to charge on the wire. The polarization current is then

$$\vec{J}_s(\rho, z, \phi) \approx \frac{-(\tilde{\epsilon}_2 - \tilde{\epsilon}_1)}{2\pi\tilde{\epsilon}_2\rho} I'(z)\hat{\rho}, \quad a \leq \rho \leq b.$$

In the thin wire approximation of the electric field integral equation, the field due to \vec{J}_s is needed on the wire axis. For a straight segment of insulated wire with $|k_1 b| \ll 1$ the field on the axis due to \vec{J}_s can be approximated, as shown in [18], in terms of the second derivative of the axial current as

$$\hat{z} \cdot \vec{E}_s(z) \approx \frac{j(\bar{\epsilon}_2 - \bar{\epsilon}_1)}{2\pi\omega\epsilon_0\bar{\epsilon}_1\bar{\epsilon}_2} I''(z) \ln(b/a) = \frac{-jk_s^2(\bar{\epsilon}_2 - \bar{\epsilon}_1)}{2\pi\omega\epsilon_0\bar{\epsilon}_1\bar{\epsilon}_2} C_1 \ln(b/a),$$

where the final result is for the current distribution of Eq. (3). This field is included in the total, axial electric field in the thin-wire integral equation.

The field on the outside of the sheath due to the polarization currents in the sheath is shown in [18] to be less than this axial field by a factor of about $k_1^2(b^2 - a^2)/4$. Hence this field can be neglected relative to the axial field for thin wires, so that the mutual interaction of segments involves only the field due to the axial wire current. The model for the sheath thus reduces to a simple form comparable in complexity to the model for a wire with finite conductivity, in that the sheath modifies the field only on its own segment. This result is particularly convenient in the case of a wire near the air-ground interface since no addition is needed to the present evaluation of the effect of the interface on the field due to the axial wire current.

One other change was needed in NEC to model insulated wires embedded in earth or water. In this case the current tends to have a sinusoidal form, as for a bare wire, but the wave number may be orders of magnitude less than for a bare wire in the medium. Use of k_s in Eq. (3) equal to the wave number in the surrounding medium can then result in very slow convergence of the solution. Fortunately, a simple approximation for the wave number on a buried insulated wire is available from the theory of coaxial transmission lines [20]. For a perfectly conducting wire, this solution predicts a wave number of

$$k_L \approx k_2 \left[1 + \frac{H_0^{(2)}(k_1 b)}{k_1 b \ln(b/a) H_1^{(2)}(k_1 b)} \right]^{1/2} \quad (7)$$

where k_1 is the wave number in the infinite medium and k_2 is the wave number in the insulating material. $H_0^{(2)}$ and $H_1^{(2)}$ are Hankel functions of order 0 and 1. This approximation should be valid for $|k_1| \gg |k_2|$. Hence, in the NEC solution, k_s is set equal to k_L when $|k_1/k_2|$ is greater than 2. For smaller $|k_1/k_2|$, k_s is set equal to k_1 . The accuracy of this estimate of k_s , within a factor of 2 or so, is not important since the solution is relatively insensitive to k_s . With a dense outer medium, the use of k_L in the current expansion greatly improves the convergence of the solution.

Allowing k_s different from k_1 increases the complexity of the field equations, however. The fields due to $\sin k_s s$ and $\cos k_s s$ current distributions, which are given by simple closed-form expressions when $k_s = k_1$, involve additional exponential integrals when $k_s \neq k_1$. The integrals were evaluated in NEC-31 by numerical integration after subtracting integrable functions to remove the peaked behavior of the integrand. Evaluation of these integrals can be difficult since, while $k_s \delta$ must be small, $k_1 \delta$ may be large when $|k_1| \gg |k_2|$. The present evaluation appears to be accurate, but more work in this area could reduce the computation time.

To validate the insulated wire model in NEC, results were compared with published measurements for insulated antennas in air and water and with the solution for the propagation constant on an infinite insulated wire. The measured input admittance of a dipole antenna with dielectric insulation in air was included in [19]. NEC results are compared with these measurements and with the results of Richmond and Newman's code in Fig. 14. The two computer codes are in close agreement for the dipole antenna, with a difference evident only toward antiresonance, as often occurs with method-of-moments codes. Comparison with NEC results for a bare dipole showed that the resonant length was reduced from $0.47\lambda_0$ to $0.42\lambda_0$ by addition of the sheath. A similar comparison for a square loop is included in [18].

A valuable check on the NEC results is provided by the analytic solution for the propagation constant on an infinite insulated wire. A rigorous solution to this problem was developed by J. R. Wait [21] who solved the boundary value problem for a horizontal insulated wire buried in the ground below an interface with air. The wire is at depth d below the air-ground interface in a medium where the

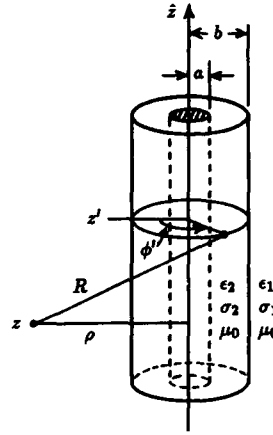


Fig. 13. Coordinates for evaluation of the field due to polarization currents in an insulating sheath on a wire.

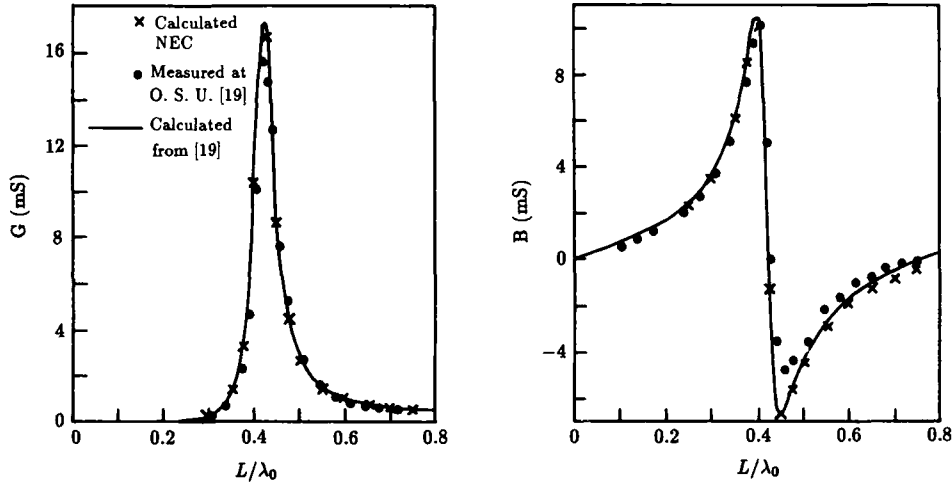


Fig. 14. Input admittance of an insulated dipole antenna in air with length L , radius $a = L/640$ and sheath radius $b = 5.84a$. The relative permittivity of the sheath is 2.3 and the conductivity is zero.

wave number is k_1 in the present notation. Assuming a current $I_0 \exp(-\Gamma x)$, Wait matched boundary conditions at the air-ground interface, the outer boundary of the sheath and on the wire, assuming that $d \gg b$ so that azimuthal variation of the field around the wire can be neglected, to derive an equation for Γ . An explicit solution for Γ is obtained in [21] when $|k_1| \gg |k_2|$. We chose here to solve Eq. (11) in [21] with a numerical root finder to ensure good accuracy over a wide range of parameters.

The transmission line model used in [20] provides another solution for the propagation constant on a buried insulated wire, which is repeated here as Eq. (7). An extension of this model to include an interface was proposed in [22] assuming an image of the wire reflected in the interface with a reflection coefficient of one. The result for a wire at depth d is

$$k_L = -j\Gamma \approx k_2 \left[1 + \frac{H_0^{(2)}(k_1 b) + H_0^{(2)}(2k_1 d)}{k_1 b \ln(b/a) H_1^{(2)}(k_1 b)} \right]^{1/2} \quad (8)$$

As with Eq. (7) it is required that $|k_1| \gg |k_2|$ in this result.

To determine the propagation constant from the NEC solution, an insulated wire several wavelengths long was modeled with a source approximately $\lambda/4$ from one end and a load approximately $\lambda/4$ from the other end, where λ should be the wavelength of the propagating wave on the wire. The position and size of the terminating load were adjusted to achieve a reasonably small standing wave. This required some trial and error since the value of λ and the optimum load value were not known in advance. The phase constant β_L for the current was then determined from the phase shift divided by distance and the attenuation constant α_L was determined from a linear match to the log of the magnitude of current. The source and termination regions were excluded in these calculations.

The propagation constants determined from the NEC-3I solution, Eq. (7) or Eq. (8) and the solution from [21] are compared for a number of cases in [18]. For an infinite medium the wire radius a was varied with b/a fixed. A typical result for a wire with air insulation in conducting water is shown in Fig. 15. Wait's solution from [21] should be the most accurate of these results over the entire range of a/λ_0 . The NEC solution is seen to be in good agreement with the result from [21] for small a/λ_0 , diverging for larger values. The result of Eq. (7) is more accurate for large a/λ_0 . Other results showed, as expected, that Eq. (7) becomes more accurate with increasing $|\epsilon_1|$. It is less accurate for small b/a and small $k_1 b$. Hence the NEC solution and Eq. (7) are often complementary in their ranges of accuracy. The value of α_L is more difficult to determine accurately than that of β_L for either Eq. (7) or NEC.

The input admittance computed by NEC-3I for an insulated monopole in conducting water with $\epsilon_1 = 80 - j0.197$ is compared in Fig. 16 with measurements made by S. Mishra and included in [20] and [23]. As described in [23], the monopoles were constructed from coaxial cable and were measured in lake

water at 300 MHz. Agreement is generally good in this case, with the NEC-3I results differing mainly in the height of the peak at resonance. Other cases included in [18] show increasing error in the peak heights for NEC-3I as b/a is increased, but good agreement in the resonant frequency. This observation confirms the results for the propagation constant on this monopole, for which $a = 3.175(10^{-3})\lambda_0$ and $k_1 b = 0.45$, since α_L from the NEC-3I solution is starting to diverge to large values while β_L is still reasonably accurate. Eq. (7) is accurate in this case but loses accuracy for smaller a/λ_0 . Hence for smaller wire and sheath radii the NEC-3I solution should become more accurate and the transmission line model less accurate.

The propagation constant on an insulated wire near an interface as determined by NEC-3I was compared with that predicted by J. R. Wait's solution in [21] and with Eq. (8). Results from NEC-3I using the Sommerfeld integral option are compared in Fig. 17 with the solution from [21]. The latter, since boundary conditions are accurately satisfied at the interface, should be equivalent to the Sommerfeld integral solution. The results for β_L are seen to be in very good agreement, while α_L shows some difference for small d but generally good agreement. The maximum error in α_L is about 5 percent.

NEC-3I results using the reflection coefficient approximation are compared in Fig. 18 with the results of Eq. (8) which also is based on a reflection coefficient treatment of the interface. The two results are in reasonably good agreement but differ substantially from the presumably accurate results in Fig. 17. Hence it appears that the plane-wave reflection coefficient is not suitable for modeling a buried wire. Although the reflection coefficient is close to one for a dense ground, the cancellation in the integral for the field appears to require a more accurate treatment.

The comparisons with measurements and other analytic results indicate that the accuracy of the NEC model for an insulated wire in a conducting medium is mainly dependent on $|k_1 b|$. The error in α_L determined by NEC reaches ten percent when $|k_1 b|$ is approximately 0.15, while the solution for β_L , reaches ten percent error at about $|k_1 b| = 0.7$.

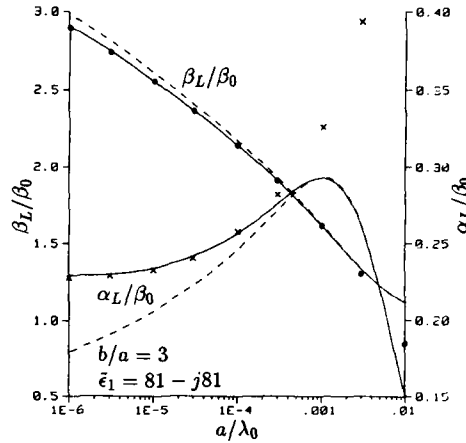


Fig. 15. Propagation constant $\Gamma = \alpha_L + j\beta_L$ versus wire radius a on a wire with air insulation in water with complex relative permittivity ϵ_1 for $b/a = 3$ and $\beta_0 = 2\pi/\lambda_0$. Results of the NEC model (• x) and Eq. (7) (---) are compared with the more accurate solution from ref. [21] (—).

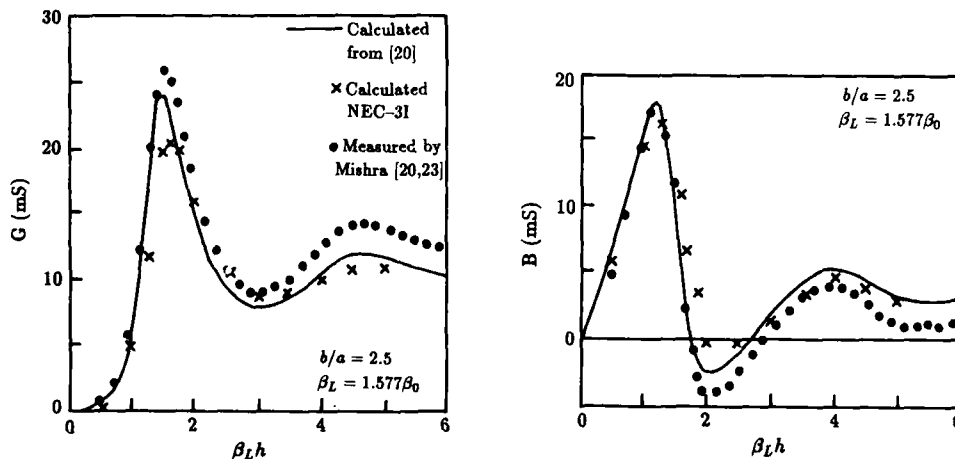


Fig. 16. Input admittance of a monopole antenna with air insulation in water with $\epsilon_1 = 80 - j0.197$ and $a = 3.175(10^{-3})\lambda_0$. The monopole height is h and β_L is the real part of k_L from Eq. (7).

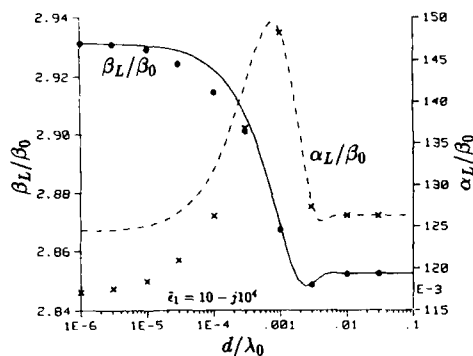


Fig. 17. Propagation constant on an insulated wire at depth d below an interface computed with the Sommerfeld integral option in NEC (• x) and solution from [21] (—, - - -). The upper half space is air, the complex relative permittivity of the medium surrounding the wire is $\epsilon_1 = 10 - j10^4$. Wire radius is $a = 2.38(10^{-7})\lambda_0$ with $b/a = 3$ and $\beta_0 = 2\pi/\lambda_0$.

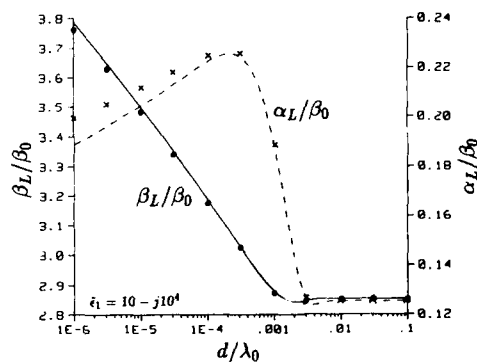


Fig. 18. Propagation constant on an insulated wire at depth d below an interface from the reflection coefficient option in NEC (• x) and Eq. (8) (—, - - -). Both results are incorrect as a result of using the reflection coefficient for the interface. Model parameters are the same as for Fig. 17.

4.5 A Hybrid of NEC-MoM and GTD

Electrically large structures, in particular those composed of surfaces, are often modeled by means of the Geometrical Theory of Diffraction (GTD), since the moment method is limited by computation time and storage. A general purpose modeling code based on GTD/UTD has been developed at the Ohio State University [24], and is known as NEC-BSC. Development of this code is continuing, with a recent version described in [25]. It can model structures composed of combinations of conducting plates, multiple elliptic cylinders, multiple sectioned cone frustums, composite finite ellipsoids, and has a limited dielectric plate capability.

Since resonant sized antennas are often affected by large conducting bodies, either through near-zone interaction or distortion of the radiation patterns, a combination of the moment method and GTD/UTD approach is suggested. This hybrid technique was demonstrated by Thiele and Newhouse [26]. A code combining NEC-3 and NEC-BSC has been developed to test the practicality of the hybrid on a larger scale [27]. The main program of NEC-BSC was converted to a subroutine that can be called in three modes: 1) to read input data for the GTD model; 2) to evaluate the field at a point due to an arbitrarily oriented point current source; and 3) to evaluate the field in a given direction due to a point source. The moment-method code NEC-HYBRID calls NEC-BSC in mode 2 in filling the interaction matrix, and in mode 3 in computing the radiation pattern. Wires can be located anywhere permitted by the GTD restriction that distance from an edge be greater than about $\lambda/4$. NEC wire models may connect to the ground plane, since the full near field of the image is computed by NEC, but not to other surfaces in the GTD model for which only the $1/R$ field is evaluated.

Results of the NEC-HYBRID code for a monopole on a 90 degree wedge are shown in Fig. 19. The difference in input resistance from that of the monopole on an infinite ground plane, R_∞ , is shown. The value of R_∞ determined by NEC-3 was 50.0 ohms, while for Thiele and Newhouse it was 53.3 ohms. The difference in the radiation patterns appears to be due to inaccuracy in the early results from [26].

Other auxiliary field evaluation routines could be adapted to link with the modified NEC-3 in place of NEC-BSC. For example, a simple wedge diffraction code has been used for fast evaluation of a wire on an infinite wedge. When the wedge in Fig. 19 was modeled with two large but finite plates in NEC-BSC, the time to fill the moment-method matrix was 15 times that for the monopole on an infinite ground plane. With the wedge diffraction code, the factor was only 2.

The hybrid approach could be extended to permit modeling the effect of terrain features on antenna radiation. Such features as hills or cliffs often have little effect on the antenna current or input impedance, affecting only the radiation pattern. At high frequencies, when only the space wave is of interest, terrain effects have been modeled by knife-edge diffraction, developed in the classic paper of Schelleng et al. [28] and later work, or by wedge diffraction modified for conductor loss [29]. At lower frequencies, propagation over terrain varying in height or conductivity has been modeled by solution

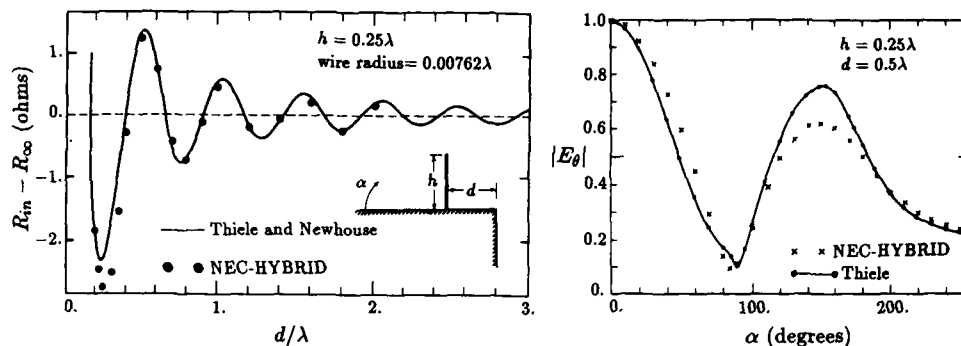


Fig. 19. Input resistance and radiation pattern of a monopole on a 90° wedge. NEC-HYBRID results are compared with those of Thiele and Newhouse [26].

of an integral equation [30]. Such propagation models can either be combined with the NEC code or used alone with input of currents from the NEC solution.

5. VALIDATION AND APPLICATIONS

Validation of a code such as NEC is an essential part of code development, and also must be a continuing consideration as the code is used in new applications. Errors, as discussed in [31], can be separated into physical modeling errors, which result from the approximations needed to represent a real structure in a form compatible with the numerical model; and numerical modeling errors, which result from approximations in the solution for the idealized model. Physical modeling errors generally must be assessed by the code user, based on understanding of physical principles or measurements. Numerical modeling errors can be evaluated by comparisons with measurements or independent analytical or numerical models, or often from self-consistency checks within the code.

Perhaps the most desirable form of validation is comparison with accurate measured results. However, measurement errors and uncertainties, and the difficulty of varying parameters in a physical model can limit the usefulness of this approach. Comparisons with independent analytical or numerical models also provide useful validation. An advantage of the latter approach is the ease of comparing over a wide range of parameters and the absence of random errors. However, one must be careful of approximations shared by the two models.

Self-consistency checks within a code such as NEC also are useful. An antenna result with negative input resistance is quickly recognized as indicating trouble. The balance of input and radiated power provides a test that is necessary, although not sufficient, to establish solution accuracy. The use of this test, in the form of average gain, was demonstrated in Table 1. Reciprocity of transmitting and receiving patterns or bistatic scattering provides a similar check. Although reciprocity may be guaranteed in a Galerkin method-of-moments solution, this is not the case in NEC. A valuable, although more computationally demanding, test is to evaluate the electric field along the surface of wires. The field will always be small at the match points if the matrix equation is being solved correctly. Usually errors in the field evaluation, basis functions or poor convergence in the model will be revealed by increased field between the match points.

Some validation tests performed with NEC are shown in the following figures. The radiation pattern of a cylinder with two attached wires in Fig. 20 demonstrates the hybrid EFIE-MFIE capability. The cylinder height is 22 cm, the diameter 20 cm, and the wire dimensions are shown on the figure. Wire a was driven by a voltage source

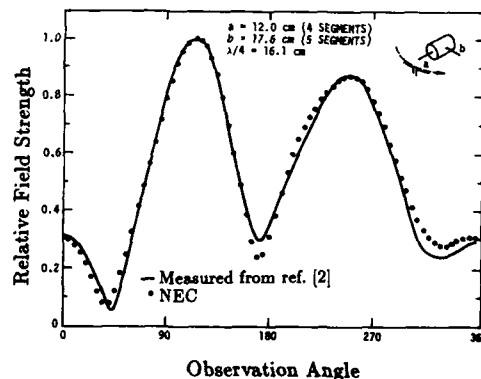


Fig. 20. Radiation Pattern of a cylinder with attached wires, with wire a excited.

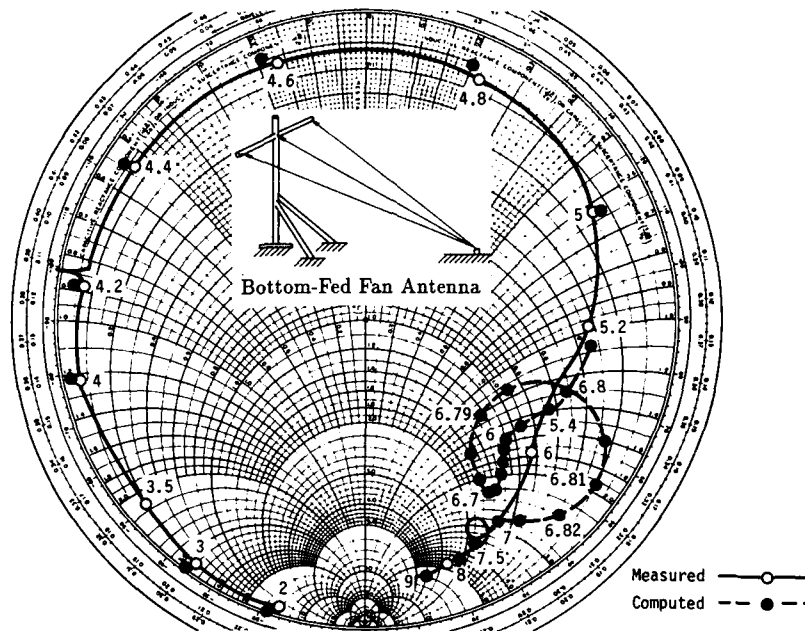


Fig. 21. Input impedance of a bottom-fed fan antenna. NEC results are compared with measurements from the Naval Ocean Systems Center on a 1/24 scale model. Frequencies are shown in MHz.

at its base, while wire *b* was connected directly to the cylinder. The measured pattern was obtained by Albertsen et al. [2] who presented similar numerical results. NEC results for a higher density of patches than was used for Fig. 20 were in worse agreement with the measurements, indicating possible problems in convergence of the MFIE solution for a surface with square edges.

The computed input impedance of a bottom-fed fan antenna is compared with measurements in Fig. 21. This result from [32] is one of several validation exercises performed for antennas used on ships [33,34]. The results for the fan are generally in good agreement from 2 to 9 MHz. However, the NEC results show a loop on the Smith chart from 6.7 to 7 MHz that does not correspond to the smaller loop in the measurements. Examination of the currents showed that this behavior was due to a transmission line-mode resonance in the wires of the fan. Such a phenomenon can be very sensitive to the exact dimensions of the structure and conductivity of the wire and insulators. The NEC model assumed perfectly conducting wires and an air gap for the insulators which may account for this difference.

The results in Fig. 21 were obtained with the three wires of the fan driven by three separate voltage sources against the ground plane. An alternate model would be to connect the three wires to a short vertical segment on which the source is located. Results in [32] for the single-source model were in worse agreement with measurements than those in Fig. 21, however. They showed a frequency shift of up to ten percent at the lower frequencies and negative input resistance over part of the transmission-line resonance around 7 MHz. However, when the single-segment source is run in the new code NEC-4X, described in Section 4.3, the results are close to those in Fig. 21. Hence the treatment of charge at the junction is the apparent source of error in NEC-3.

Verification of accuracy is a particular concern when modeling antennas over a finitely conducting ground due to the complexity of the ground model. As discussed in [8], internal checks such as integrating the total radiated power for a dielectric ground, and numerical checks on boundary conditions along wires or across the interface were quite useful for validation. NEC-3 was in good agreement with an independent moment-method code for a vertical cylinder passing through the interface with no *a priori* condition on charge density [35].

The wave number for current on a horizontal wire over the ground provides a test that is sensitive to the accuracy of the ground model and can conveniently be compared with analytic solutions. Such a comparison for a buried insulated wire was demonstrated in Figs. 17 and 18. For a bare wire above

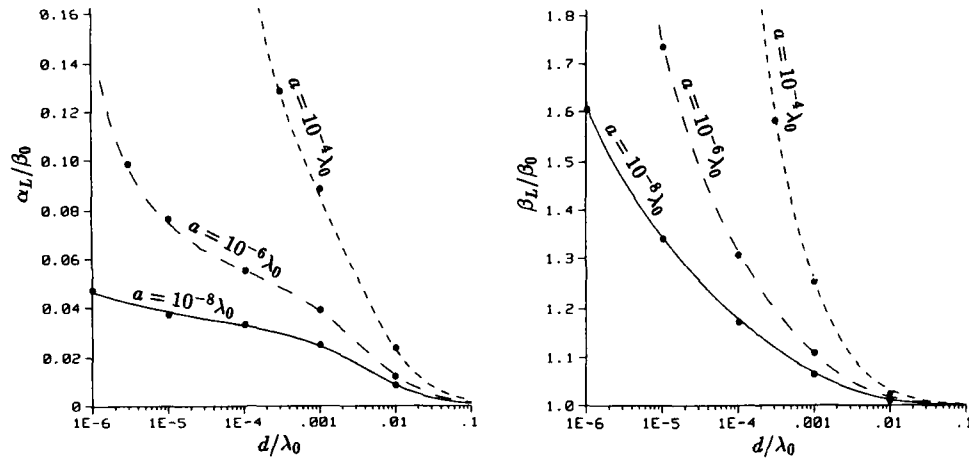


Fig. 22. Wave number $k_L = \beta_L - j\alpha_L$ on a horizontal wire with radius a at height d above a ground plane. NEC results (points) are compared with results of the eccentrically insulated transmission line model (lines) from [36].

ground, an approximation for the wave number was derived in [36] from the equations for a coaxial transmission line with lossy outer conductor and eccentrically located inner conductor. The outer conductor radius is allowed to go to infinity so that the transmission line becomes a wire at height d above a flat ground plane. Approximations in this model require that the wave number in the ground be much greater than that in the medium surrounding the wire. The propagation constant from this formula has been compared with measurements for a wire over water with good agreement [37].

The wave number k_L determined from NEC is compared with that from the transmission line approximation in Fig. 22 for a ground with $\bar{\epsilon}_g = 10 - j1000$ and for wire radii of $10^{-8}\lambda_0$, $10^{-6}\lambda_0$ and $10^{-4}\lambda_0$. This complex relative permittivity would correspond to a reasonably good ground ($\sigma = 0.01$ S/m) at about 200 kHz. The agreement is seen to be very good for this case. For lower ground conductivity the transmission line approximation would lose accuracy, while NEC would be expected to remain accurate.

The results in Fig. 22 represent a test of the accuracy of the field evaluation very near the interface. Another critical factor in the ground model is the treatment of charge on wires passing through the interface or junctions near the interface. This was a particular concern in a special version of NEC-3 (NEC-GS) developed to model a monopole on a uniform radial-wire ground screen. By taking full advantage of the physical symmetry and excitation symmetry, NEC-GS can model large radial-wire screens in much less time and storage than would be required by NEC-3. Little accurate experimental data is available for comparison, however, so validation has relied on self-consistency checks such as convergence and power balance. Also, a comparison with an approximation developed by Wait and Pope [38] has helped to validate the NEC-GS results.

In the NEC model for a ground screen, the junction is located exactly on the interface with the first segment on each radial sloping downward to the desired depth of the screen. One question in this model was the treatment of charge at the junction of the monopole and the radial wires, where many wires meet with narrow angles of separation. In the current expansion, the derivative of current on the monopole just above the interface (I'_+) and on the radials below ground (I'_-) is made to satisfy the condition $I'_-/I'_+ = \bar{\epsilon}_g$, where $\bar{\epsilon}_g = \epsilon_- - j\sigma_-/\omega\epsilon_0$ is the complex relative permittivity of the ground. This condition was derived for continuity of radial electric field of a vertical wire penetrating the interface [8], so its application to the ground screen could be questioned.

Results for buried ground screens are generally found to be stable with varying segment lengths, however. The most difficulty has been encountered in modeling ground screens floating above the ground. In this case, since all wires are in air, the charge densities on wires at the junction are forced to be equal for equal wire radii. To obtain converged results for ground screens at small heights above the ground, it was necessary to use segment lengths at the junction on the order of the height of the screen above ground, tapering to larger segments away from the junction. This is particularly important for

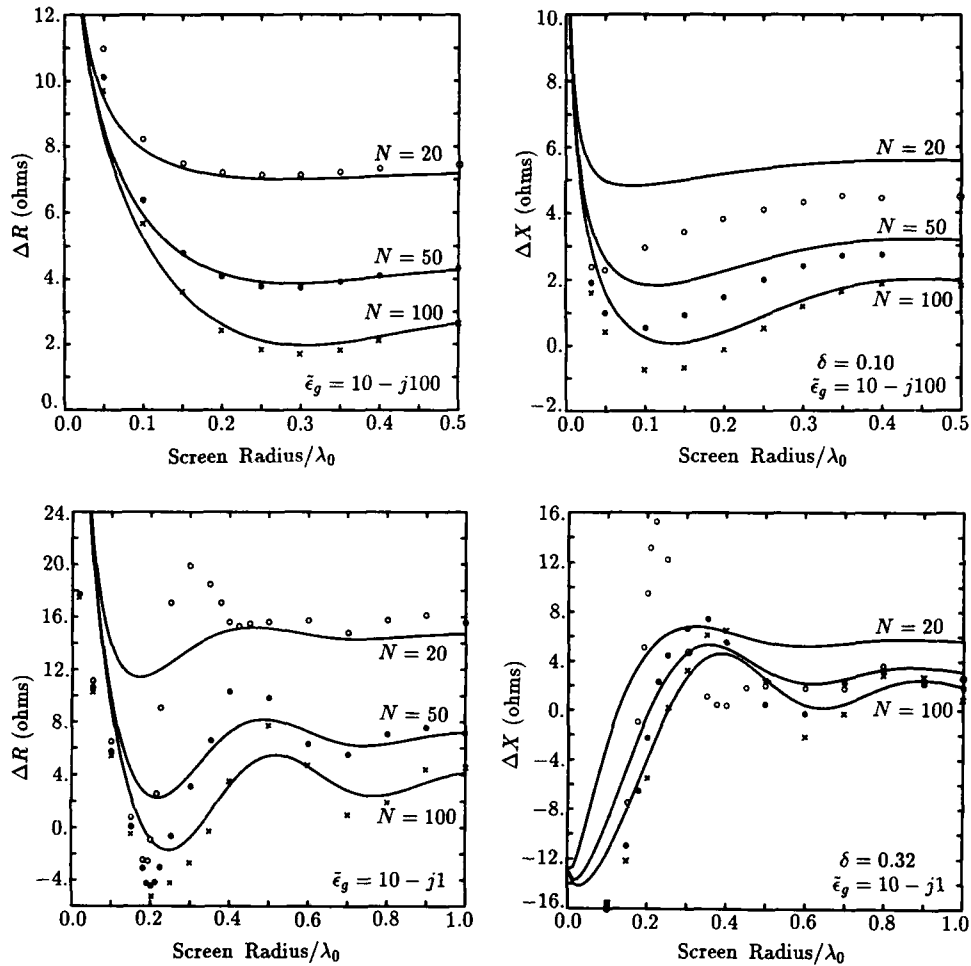


Fig. 23. Input impedance of a $\lambda/4$ monopole on a buried radial wire ground screen with N wires, showing the difference from the impedance on a perfectly conducting ground plane. NEC results (points) are compared with results from a result derived from the compensation theorem [38].

electrically small monopoles and screens. A treatment of charge similar to that described in Section 4.3, but taking the ground into account, appears necessary to obtain an optimum basis function.

As a check on the solution for buried ground screens, the computed input impedance of a monopole was compared with that predicted by an approximation developed by Wait and Pope [38]. For this approximation, a current $I_0 \cos(kz)$ is assumed on the monopole. The difference in input impedance (ΔZ) between that of the monopole on the radial-wire screen in real earth and that of the monopole on an infinite perfectly conducting ground is obtained as an integral over the surface, through application of the compensation theorem. The magnetic field over the ground surface is taken to be that on the perfect ground, and the electric field is related to it by a surface impedance for the screen and ground, or for bare ground beyond the screen. A derivation of this approximation and results are included in [39]. Wait states that this approximation should be accurate for a screen radius in wavelengths (b/λ_0) somewhat greater than $\delta = |\bar{\epsilon}_g|^{-1/2}$. Also, the number of wires in the ground screen must be sufficiently large for the surface impedance approximation to be valid.

The ΔZ computed by NEC-GS is compared with that from the compensation theorem result in Fig. 23. The input impedance computed by NEC-GS for the monopole on an infinite perfectly conducting ground was $38.6 + j22.2$ ohms. Agreement in ΔZ is generally good for b/λ_0 greater than δ . The

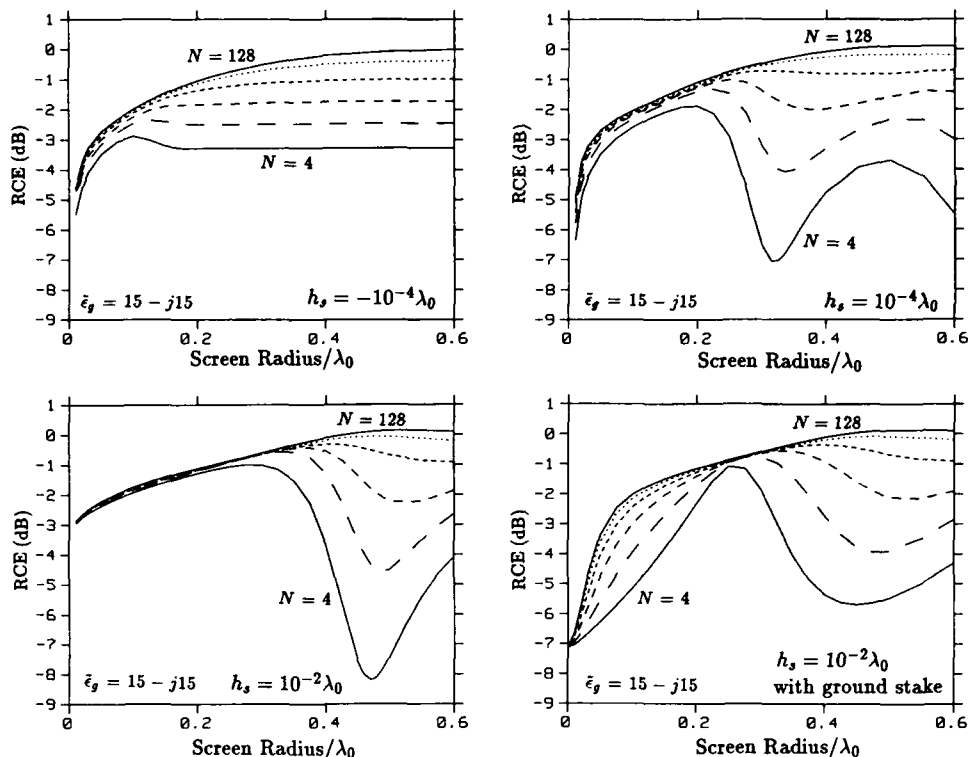


Fig. 24. Relative Communication Efficiency of a $\lambda_0/4$ monopole on a radial-wire ground screen below and above the surface of the ground. The height of the screen above the ground is h_s , and the number of radials N is 4, 8, 16, 32, 64 and 128.

larger discrepancies for 20 radials are most likely due to failure of the surface impedance approximation for a sparse screen. The discrepancies for low ground conductivity are due to standing waves on the screen wires which are not taken into account by the compensation theorem solution. This comparison supports the accuracy of the NEC model, which should remain accurate for sparse screens and low ground conductivity. Unfortunately, the compensation-theorem result fails for small screen radius where the NEC model appears most in doubt.

Some results for antennas operated over ground follow to demonstrate the usefulness of the code for parametric studies. Fig. 24 shows the Relative Communication Efficiency (RCE) of a $0.25\lambda_0$ monopole on a radial-wire ground screen buried and above the ground. RCE is defined as the ratio of power density at some receiver location due to the test antenna to that due to a reference antenna for equal input powers. The reference antenna in this case was the monopole on a buried screen with 128 radials, and the receiver location was on the ground $1000\lambda_0$ from the transmitter.

While wires exactly on the interface cannot be modeled due to the thin wire approximation, the results for screens $10^{-4}\lambda_0$ below and above the ground surface are nearly the same for small screen radius, until current attenuation limits the buried screen. Sparse screens above ground show a sharp loss of efficiency when the radials are near to a half-wavelength resonance. For a screen height of $10^{-4}\lambda_0$ the interaction with the ground shifts this resonance down to where it reduces the RCE for four radials as short as $0.25\lambda_0$. Inclusion of a ground stake under the monopole, in addition to the radial-wire screen, is sometimes recommended, either to improve antenna performance or for safety. The last result in Fig. 24, in which the stake depth was $0.02\lambda_0$, and results for other grounds show mainly detrimental effects for the ground stake. Multiple shorter stakes were more effective but still reduced efficiency, especially for small screen radius. With lower ground conductivity, the buried screen was also affected by screen resonances, as shown in Fig. 25.

The RCE of a Beverage antenna is shown in Fig. 26 for varying length and height. In this case, the reference antenna used to normalize the results was a vertical dipole with length $0.1\lambda_0$ on the interface.

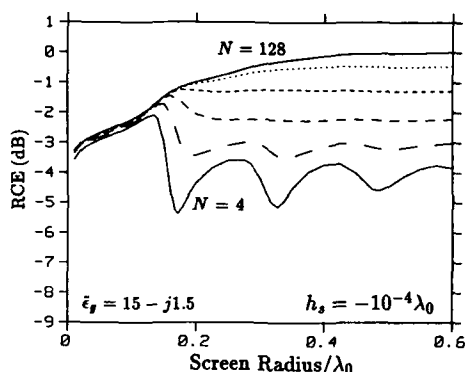


Fig. 25. Relative Communication Efficiency of a $\lambda_0/4$ monopole on a radial-wire ground screen buried by $10^{-4}\lambda_0$. The number of radials N is 4, 8, 16, 32, 64 and 128.

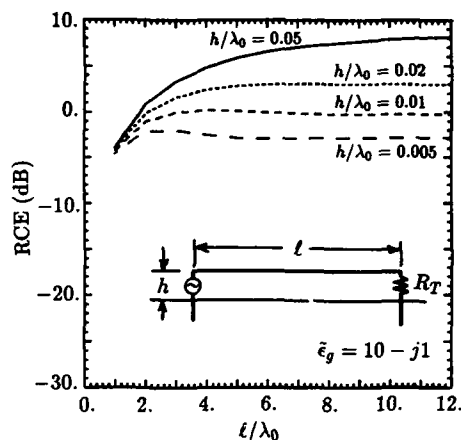


Fig. 26. Relative Communication Efficiency of a Beverage antenna for varying length, with height a parameter.

The limitation of gain due to phase lag along the wire is evident as the length is increased.

NEC has been applied to generate an antenna engineering handbook and data base [40]. This effort included basic field-antenna types: whips on ground stakes, screens and boxes; half-squares, vee beams, horizontal dipoles, inverted vee dipoles, long wires and loops. Tables of radiation patterns were generated for a range of frequencies and grounds.

6. CONCLUSION

NEC is a versatile code for modeling antennas and their environment. It has become widely used, due mainly to its convenient operation, documentation, availability and continuing support. Recent extensions to NEC include a model for insulated wires and improved accuracy in modeling electrically small antennas and discontinuities in wire radius. Valuable extensions for future consideration include the addition of an EFIE model for open surfaces such as plates and a model for antennas in or over stratified ground. A longer range goal is to incorporate expert-systems technology to simplify the process of building models that conform to the modeling guidelines and to verify that the code is producing correct results.

The utility of NEC has been enhanced by the development of an interactive graphics utility (IGUANA) [41] for developing the model description and examining results. IGUANA supports both NEC and the moment-method code MININEC [13] for use on personal computers. A package NEEDS-1.0 is now available from the Applied Computational Electromagnetics Society (ACES) that includes NEC-2 for an IBM PC as well as IGUANA, MININEC and the GRAPS graphics package.

NEC is written in Fortran and consists of about 10,000 lines of code. Versions are available for DEC/VAX, CDC, CRAY and IBM, and it has been converted to a number of other computers. Of the versions currently being released, NEC-2 has no distribution restrictions, while NEC-3 and other versions in development require authorization from the U. S. Department of Defense.

The development of NEC has been sponsored by the U. S. Army ISEC (J. McDonald, L. Corrington) and CECOM (P. Cunningham) and the U. S. Navy NOSC (J. Logan, S. T. Li). The work to model insulated wires was partially sponsored by the U. S. Air Force, RADC/EECT and RADC/DCCL.

REFERENCES

- [1] G. J. Burke and A. J. Poggio, *Numerical Electromagnetics Code (NEC) - Method of Moments*, Lawrence Livermore National Laboratory, Rept. UCID-18834, January 1981.
- [2] N. C. Albertsen, J. E. Hansen and N. E. Jensen, *Computation of Spacecraft Antenna Radiation Patterns*, The Technical University of Denmark, 1972.

- [3] A. J. Poggio and E. K. Miller, "Integral Equation Solutions of Three-Dimensional Scattering Problems," Chapt. IV in *Computer Techniques for Electromagnetics*, edited by R. Mittra, Pergamon Press NY, 1973.
- [4] A. J. Poggio and R. W. Adams, *Approximations for Terms Related to the Kernel in Thin-Wire Integral Equations*, Lawrence Livermore National Lab. Rept. UCRL-51985, December 1975.
- [5] R. F. Harrington, *Field Computation by Moment Methods*, McGraw Hill Book Company, New York, 1961.
- [6] Y. S. Yeh and K. K. Mei, "Theory of Conical Equiangular Spiral Antennas," Part I - Numerical Techniques, *IEEE Trans. Antennas and Propagation*, Vol. AP-15, p. 634, 1967.
- [7] T. T. Wu and R. W. P. King, "The Tapered Antenna and its Application to the Junction Problem for Thin Wires," *IEEE Trans. Antennas and Propagation*, Vol. AP-24, pp. 42-45, 1976.
- [8] G. J. Burke and E. K. Miller, "Modeling Antennas Near to and Penetrating a Lossy Interface," *IEEE Trans. Antennas and Propagation*, Vol. AP-32, No. 10, pp. 1040-1049, 1984.
- [9] G. J. Burke, "Test of a Coaxial-Line Source Model for NEC," Proceedings of the 3rd Annual Review of Progress in Applied Computational Electromagnetics, Monterey, CA, March 24-26, 1987.
- [10] G. J. Burke, *Enhancements and Limitations of the Code NEC for Modeling Electrically Small Antennas*, Lawrence Livermore National Laboratory, Rept. UCID-20970, January 1987.
- [11] G. J. Burke, *Treatment of Small Wire Loops in the Method of Moments Code NEC*, Lawrence Livermore National Laboratory, Rept. UCID-21196, October 1987.
- [12] D. R. Wilton, University of Houston, private communication.
- [13] J. C. Logan and J. W. Rockway, *The New MININEC (Version 3): A Mini-Numerical Electromagnetic Code*, Naval Ocean Systems Center TD 938, September 1986.
- [14] A. W. Glisson and D. R. Wilton, *Numerical Procedures for Handling Stepped-Radius Wire Junctions*, Department of Electrical Engineering, University of Mississippi, March 1979.
- [15] A. C. Ludwig, "Wire Grid Modeling of Surfaces," *IEEE Trans. Antennas and Propagation*, Vol. AP-35, pp. 1045-1048, 1987.
- [16] P. C. Waterman, "Matrix Formulation of Electromagnetic Scattering," *Proceedings of the IEEE*, Vol. 53, pp. 805-812, August 1965.
- [17] J. P. Desmettre, Commissariat a L'Energie Atomique, France, Private Communication.
- [18] G. J. Burke, *A Model for Insulated Wires in the Method of Moments Code NEC*, Lawrence Livermore National Laboratory, Rept. UCID-21301, January 1988.
- [19] J. H. Richmond and E. H. Newman, "Dielectric Coated Wire Antennas," *Radio Science*, Vol. 11, no. 1, pp. 13-20, Jan. 1976.
- [20] R. W. P. King and G. S. Smith, *Antennas in Matter*, The MIT Press, Cambridge, MA, 1981.
- [21] J. R. Wait, "Electromagnetic Wave Propagation Along a Buried Insulated Wire," *Canadian J. of Phys.*, Vol. 50, pp. 2402-2409, 1972.
- [22] R. W. P. King, "Antennas in Material Media Near Boundaries with Application to Communication and Geophysical Exploration, Part II: The Terminated Insulated Antenna," *IEEE Trans. Antennas and Propagation*, Vol. AP-34, no. 4, pp. 490-496, April 1986.
- [23] R. W. P. King, S. R. Mishra, K. M. Lee and G. S. Smith, "The Insulated Monopole: Admittance and Junction Effects", *IEEE Trans. Antennas and Propagation*, Vol. AP-23, no. 2, pp. 172-177, March 1975.
- [24] R. J. Marhefka and W. D. Burnside, "Numerical Electromagnetics Code - Basic Scattering Code, NEC-BSC (Version 2), Part I: User's Manual," Technical Report 712242-14, December 1982, The Ohio State University ElectroScience Laboratory, Department of Electrical Engineering; prepared under Contract No. N00123-79-C-1469 for Naval Regional Contracting Office.

- [25] R. J. Marhefka, "Development and Validation of a New Version of the NEC-BSC," *Proceedings of the 4th Annual Review of Progress in Applied Computational Electromagnetics*, Monterey, CA, March 22-24, 1988.
- [26] G. A. Thiele and T. H. Newhouse, "A Hybrid Technique for Combining Moment Methods with the Geometrical Theory of Diffraction," *IEEE Trans. Antennas and Propagation*, Vol. AP-23, pp. 62-69, January 1975.
- [27] G. J. Burke, *NEC-HYBRID User's Guide*, Lawrence Livermore National Laboratory, Rept. UCID-20959, 1987.
- [28] J. C. Schelleng, C. R. Burrows and E. B. Ferrell, "Ultra-Short-Wave Propagation," *Proceedings of the IRE*, Vol. 21, No. 3, March, 1933.
- [29] R. J. Luebbers, "Propagation Prediction for Hilly Terrain Using GTD Wedge Diffraction," *IEEE Trans. Antennas and Propagation*, Vol. AP-32, No. 9, Sept. 1984.
- [30] R. H. Ott and L. A. Berry, "An Alternative Integral Equation for Propagation over Irregular Terrain," *Radio Science*, Vol. 5, No. 5, pp. 767-771, May 1970.
- [31] E. K. Miller, "A Selective Survey of Computational Electromagnetics," *IEEE Trans. Antennas and Propagation*, Vol. 36, No. 9, pp. 1281-1305, Sept. 1988.
- [32] G. J. Burke and A. J. Poggio, *Computer Analysis of the Bottom-Fed Fan Antenna*, Lawrence Livermore National Laboratory, Rept. UCRL-52109, August 1976.
- [33] G. J. Burke and A. J. Poggio, *Computer Analysis of the Twin-Whip Antenna*, Lawrence Livermore National Laboratory, Rept. UCRL-52080, June 1976.
- [34] F. J. Deadrick, G. J. Burke and A. J. Poggio, *Computer Analysis of the Trussed-Whip and Discone-Cage Antennas*, Lawrence Livermore National Laboratory, Rept. UCRL-52201, January 1976.
- [35] W. A. Johnson, "Analysis of a vertical, tubular cylinder which penetrates an air-dielectric interface and which is excited by an azimuthally symmetric source," *Radio Science*, Vol. 18, No. 6, pp. 1273-1281, 1983.
- [36] R. W. P. King, T. T. Wu and L. C. Shen, "The horizontal wire antenna over a conducting or dielectric half space: current and admittance," *Radio Sci.*, Vol. 9, No. 7, pp. 701-709, 1974.
- [37] R. M. Sorbello, R. W. P. King, K. M. Lee, L. C. Shen and T. T. Wu, "The horizontal wire antenna over a dissipative half-space: generalized formula and measurements," *IEEE Trans. Antennas and Propagation*, Vol. AP-25, No. 6, 1977.
- [38] J. R. Wait and W. A. Pope, "The Characteristics of a Vertical Antenna with a Radial Conductor Ground System," *Appl. Sci. Res.*, Sec. B, Vol. 4, pp. 177-195, 1954.
- [39] J. R. Wait, "Characteristics of Antennas Over Lossy Earth," Chapter 23 in *Antenna Theory: Part II*, ed. R. E. Collin and F. J. Zucker, McGraw Hill Book Co., Inc., New York, 1969.
- [40] J. K. Breakall, E. E. Domning and A. M. Christman, *Antenna Engineering Handbook*, Lawrence Livermore National Laboratory, Rept. UCID-21271, 1987.
- [41] J. Strauch and S. Thompson, *Interactive Graphics Utility for Army NEC Automation (IGUANA)*, Naval Ocean Systems Center, Rept. CR 308, 1985.

SOME APPLICATIONS OF MODEL-BASED PARAMETER ESTIMATION IN COMPUTATIONAL ELECTROMAGNETICS

E. K. Miller
General Research Corporation
5383 Hollister Avenue
Santa Barbara, CA 93111
G. J. Burke
Lawrence Livermore National Laboratory
PO Box 808
Livermore, CA 94550

ABSTRACT

All of electromagnetics, whether measurement, analysis, or computation, may be regarded as activities of information acquisition, processing, and presentation. It is relevant to mention the information-intensive nature of electromagnetics because this property can provide a useful, unifying perspective from which might be developed efficiency improvements in these various areas. If, for example, redundant information can be reduced or avoided in numerical calculations, the result should be a possibly substantial decrease in the amount of computer time that is needed for computer modeling.

In this lecture, we consider the use of a generalized signal-processing approach called "model-based parameter estimation" (MBPE) for making various improvements in either the efficiency of electromagnetic computer modeling or the efficiency of representing electromagnetic observables. One is to increase the efficiency of the basic model computation itself, a typical example being that of replacing the rigorous Green's function for the antenna-interface problem by a simpler, more easily computable, approximation. Another is to make more efficient use of the results that are computed, for example, by reducing the number of samples needed to construct a transfer function over some frequency band. A third arises because the electromagnetic fields/sources in various transformed-pair domains, such as exhibited by the far fields and locations of a linear-source array, are described by exponential and pole series from which physically relevant parameters can be extracted, which are useful for such purposes as data compression and physical interpretation. The discussion below deals primarily with MBPE based on exponential- and poles-series models which yield generalized waveform- and spectral-domain response functions of various transform-pair variables. We demonstrate how sampling such observables in terms of the appropriate variable or derivatives thereof leads to a data matrix from which the model parameters can be computed. Various kinds of data and models are used to illustrate the breadth of potential applications, with an emphasis on estimating frequency responses from frequency-sampled data. Some concluding remarks are addressed to use of MBPE for the Green's-function application.

1 INTRODUCTION

One of the most commonly encountered problems in electromagnetics, and a prototype for demonstrating the benefits of MBPE applications in computational electromagnetics, is that of determining the response of an antenna, propagation path, or a scatterer over a spectrum of frequencies rather than at just one or a few isolated frequency points. This is increasingly the case as the operating bandwidths of systems increase or their transient responses are required. A representative example is presented in Figure 1.1 where the computed (using NEC) admittance of a "forked-monopole" antenna is plotted as a function of frequency, to illustrate the difficulty that "point" sampling of a sharply-peaked transfer function can impose. Unfortunately, the samples used and their spacing in frequency seem to be determined more often by the esthetic appearance of straight-line or other low-order interpolation rather than by exploiting the underlying physics of the phenomenon being modeled. The result can be that many more frequency samples are used than should be necessary from sampling requirements based on information theory, with a consequent but avoidable increase in the cost of the associated computation or measurement.

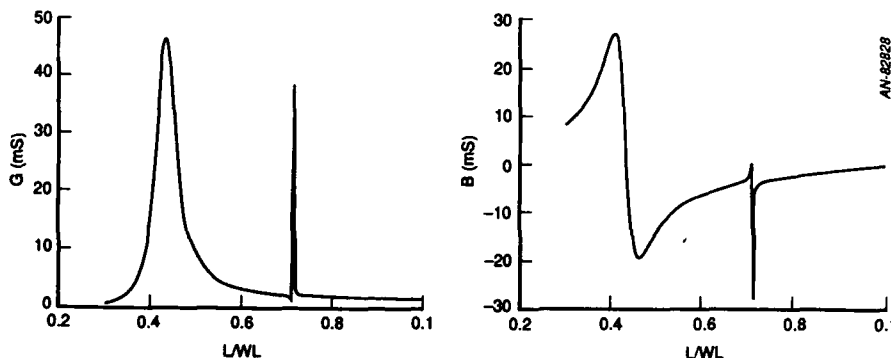


Figure 1.1. The input admittance of a forked-monopole antenna as a function of frequency. This result is representative of the kinds of transfer functions that can be encountered in electromagnetics and demonstrates the potential difficulty of adequately representing such frequency dependencies using only point sampling.

As an alternative to this interpolatory approach, perhaps best described as curve-fitting, we discuss and demonstrate techniques which incorporate a model of the expected behavior based on electromagnetic physics, which can be called "model-based parameter estimation" (MBPE). It is the physical origination of the model which takes MBPE beyond mere curve fitting, to which it might be related by the term "smart" curve fitting, as it can be appreciated that by employing "fitting functions" whose mathematical form arises from the physics of the problem at hand more efficient models will result. In the case of representing a frequency response for example, the use of polynomials or exponential functions of pure-imaginary argument (a Fourier series, for example) is not well advised because such functions are unsuited for the resonance structure commonly exhibited by electromagnetic transfer functions. Another example of this basic idea, applied in a quite different context and naturally based on different model functions, is that of replacing the Sommerfeld integrals which occur in the interface problem by simple analytical approximations whose amplitudes are determined numerically by matching them to accurately computed Sommerfeld integrals [1]. While it should be emphasized that MBPE can employ any appropriate functions as the model, many electromagnetic phenomena are well suited for representation by series of complex exponentials (in the waveform domain) or poles or more generally rational functions (in the spectral domain), for which a more thorough treatment is given by Miller [2], and which is the primary area of attention in this discussion.

To continue with the prototype problem of obtaining a wideband frequency response, we note that the most straightforward MBPE approach would be to use discrete frequency samples to develop a functional representation of the response being sought. But a second approach is also possible which takes advantage of an evidently previously unused property of the moment-method impedance and admittance matrices to obtain derivatives of the frequency response for much less computational cost than that required to obtain the original response at that same frequency. Both of these approaches utilize the mathematical property that the frequency response is well approximated by a rational function or ratio of two polynomials (the model) in which the variable is the complex frequency and for which the polynomial coefficients (the parameters) are to be found. The problem then becomes one of determining numerical values for the unknown coefficients which appear in the rational function by using whatever appropriate data might be available. As we will show, the possibility then also arises of developing a more general moment-method admittance matrix whose coefficients are themselves functions of complex frequency, leading to further interesting ramifications for obtaining and representing both spectral and temporal responses. Because exponential and pole series have so many EM applications, we outline MBPE based on each kind of model in some detail, and include representative results to demonstrate various applications. We conclude our presentation by discussing estimation applications which use other kinds of models.

2 MODEL-BASED PARAMETER ESTIMATION USING EXPONENTIALS AND POLES

2.1 The Basic Waveform (Exponential-Series) and Spectral (Pole-Series) Models

Many kinds of mathematical transformations are used in solving physical problems. One of the more useful of these, and one which is ideally suited to the exponential forms of interest here, is the Laplace transform. For the generic waveform-domain series having W exponential terms as a function of waveform-variable x given by

$$f(x) = f_p(x) + f_{np}(x) = \sum_{\alpha=1}^W R_{\alpha} \exp(s_{\alpha} x) + f_{np}(x) \quad (2.1)$$

the single-sided Laplace transform leads to the generic spectral-domain pole series as a function of transform-variable X given by

$$F(X) = F_p(X) + F_{np}(X) = \sum_{\alpha=1}^W \frac{R_{\alpha}}{(X - s_{\alpha})} + F_{np}(X) \quad (2.2)$$

where we use the terms waveform and spectral domains as the most commonly encountered physical examples given by the forms above, respectively. The subscripts "p" and "np" refer to the pole and nonpole parts of $f(x)$ and $F(X)$ respectively about which more will be said below.

We note that knowledge of the W s_{α} 's (the poles) and the W R_{α} 's (the residues) provides complete information both for $f(x)$ and $F(X)$ in the absence of the nonpole term as depicted generically in Figure 2.1. This means that rather than needing many sampled values to reconstruct a waveform or spectrum, all that needs to be stored are the two sets of numbers (generally complex) s_{α} and R_{α} . The result can be a greatly reduced number of bits required for storing such data while also providing a function for extrapolating or interpolating $f(x)$ or $F(X)$, continuous in the independent variable. A fairly wide range of EM observables have the mathematical form of Equations (2.1) and (2.2), some of which are summarized in Table 1.

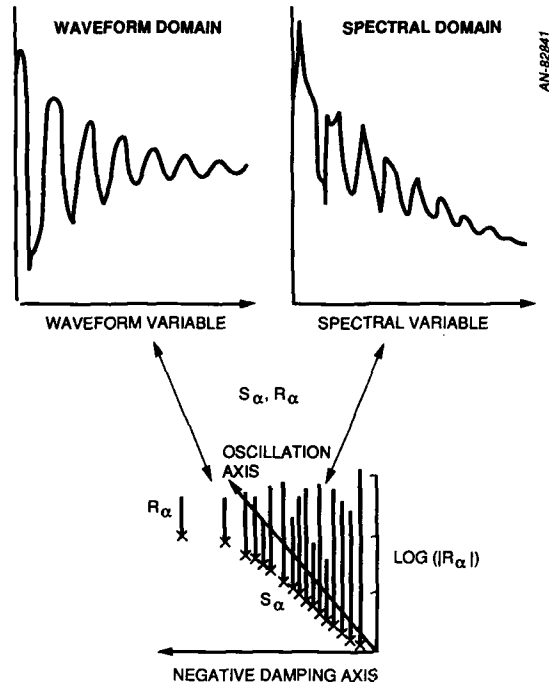


Figure 2.1. Generic waveform and pole results. Results are presented here for the backscattered field produced by a broadside-incident Gaussian-pulse plane wave on a straight wire. The waveform-domain (time) and spectral-domain (frequency) behavior are related by the poles and zeroes that in this case have been estimated from the computed time response using Equations (3.4) and (3.5).

Table 1
VARIOUS WAVEFORM- AND SPECTRAL-DOMAIN FIELD EXPRESSIONS

DOMAIN	VARIABLES		MODEL	PARAMETERS
WAVEFORM (Exponential)	x	g	$f(x)$	Residues
SPECTRAL (Pole)	X		$F(X)$	Poles
Time	t	1	$f(t) = \sum M_\alpha \exp(s_\alpha t)$	M_α = amplitude of α th mode.
Complex Frequency	s		$F(s) = \sum M_\alpha / (s - s_\alpha)$	$s_\alpha = i\omega_\alpha$, σ_α = complex-resonance frequency of α th mode.
Frequency	ω	i/c	$f(\omega) = \sum S_\alpha \exp(g\omega R_\alpha)$	S_α = amplitude of α th source.
Space	R		$F(R) = \sum S_\alpha / [g(R - R_\alpha)]$	R_α = position of α th source along line of view.
Space	x	ik	$f(x) = \sum P_\alpha \exp[gx \cos(\phi_\alpha)]$	P_α = amplitude of α th incident plane wave.
Angle	ϕ		$F(\phi) = \sum P_\alpha / [g[\cos(\phi) - \cos(\phi_\alpha)]]$	ϕ_α = incidence angle of α th plane wave with respect to line of observation (x).
Angle	ϕ	ik	$f(\phi) = \sum S_\alpha \exp[gx_\alpha \cos(\phi)]$	S_α = amplitude of α th source.
Space	x		$F(x) = \sum S_\alpha / [g(x - x_\alpha)]$	x_α = position of α th source along linear array.

2.2 Parameter Estimation

Practical implementation of Eqns. (2.1) or (2.2) requires that s_a and R_a be quantifiable from whatever analytical, computed, or measured data or information that is available. A procedure for doing this from waveform data was developed in 1795 by Prony [3,4]. Prony's method in various modified and extended forms may be recognized in many signal-processing methods in present use (see [5] for example). As outlined below, Prony's Method can be used as originally developed, for processing waveform data, and in a slightly different, but closely analogous version, for processing spectral data using in both cases what we term function sampling. Further modifications permit processing of waveform and spectral data where the data samples are derivatives with respect to the independent variable, to which we refer as derivative sampling. We derive first the waveform-domain version of Prony's method, followed by its spectral-domain counterpart for both function sampling and derivative sampling.

We note that throughout the discussion below, W denotes the number of pole terms used in the model from which the data is generated, and that P denotes the number of pole terms employed in the model from which the parameters are estimated. For valid results to be obtained, it is necessary that $P \geq W$, but since W is generally unknown *a priori*, it should also be regarded as a parameter to be determined. Also, the parameters of the generation model are indicated by ordinary letters and those of the estimation model by overscript carats, respectively. Finally, D is used to denote the number of data samples employed in the estimation process.

3 PROCESSING IN THE WAVEFORM DOMAIN

3.1 Introduction

Obtaining the pole-based parameters of an exponential-series, or waveform-domain, model is the original problem for which Prony's Method was developed. Because other variants of Prony's Method share its fundamental characteristics, it is outlined in detail below. We then present an alternate procedure which uses waveform-derivative rather than waveform-function sampling, for estimating the same parameters.

3.2 Using Waveform-Function Sampling

Determination from waveform-function sampling of the poles in Eq. (2.1) using Prony's method involves sampling $f(x)$ at D equally-spaced data points $x_i = i\delta$, $i = 0, 1, \dots, D-1$ over an interval of width $\Delta = \delta(D-1)$ from which a linear system of equations is derived in terms of an intermediate set of "linear-predictor" coefficients. Using uniform steps in the waveform variable x is necessary because the data model must be recast in polynomial form where the powers of the individual terms in the polynomial are determined by the change in observation variable between successive terms.

We first rewrite a sampled representation of Eq. (2.1) in terms of estimated parameter values as

$$f_i = f(x_i) = \sum_{n=1}^P \hat{R}_n (\hat{Z}_n)^i; \quad i = 0, 1, \dots, D-1 \quad (3.1)$$

where

$$\hat{Z}_n = e^{s_n \delta}.$$

P is the number of poles being sought, the caret denotes a parameter-value estimate, and we consider $f(x)$ to be a waveform having only pole terms where $f_{sp}(x) = 0$. We note that were $f(x)$ an electromagnetic transient response, a model having the form of Eq. (3.1) is rigorously applicable only in the "source-free" regime where the excitation which has produced the response is zero or is no longer affecting it. During the driven portion of the response, the nonpole part of Eq. (3.1) would also need to be modeled as would the time-dependent nature of the R_n 's, whose amplitudes are essentially "pumped up" by the excitation phase of the interaction.

Eq. (3.1) can be expanded as

$$\begin{aligned} \hat{R}_1 \hat{Z}_1 + \hat{R}_2 \hat{Z}_2 + \dots + \hat{R}_P \hat{Z}_P &= f_1 \\ \hat{R}_1 (\hat{Z}_1)^2 + \dots + \hat{R}_P (\hat{Z}_P)^2 &= f_2 \\ \vdots & \\ \hat{R}_1 (\hat{Z}_1)^{D-1} + \hat{R}_2 (\hat{Z}_2)^{D-1} + \dots + \hat{R}_P (\hat{Z}_P)^{D-1} &= f_{D-1} \end{aligned} \quad (3.2)$$

Now suppose the $\exp(s_n \delta)$ represent the roots of a P -th order polynomial, i.e.,

$$(3.3a)$$

$$\sum_{\alpha=1}^P \hat{a}_{\alpha} Z^{\alpha} = 0; \text{ for } Z = Z_i, i = 1, 2, \dots, P$$

Upon multiplying each row of the series of Eq. (3.2) respectively in turn by

$$\hat{a}_0, \hat{a}_1, \dots, \hat{a}_P$$

and adding the first P of these equations together, while making use of Eq. (3.3a), there can be derived

$$\hat{a}_0 f_0 + \hat{a}_1 f_1 + \dots + \hat{a}_{P-1} f_{P-1} + \hat{a}_P f_P = 0$$

If this operation is repeated D - P - 1 times by successively beginning with the second equation of Eq. series (3.2), ..., up to the D - P - 1 equation, the following D - P linear equations will result

$$\begin{bmatrix} f_0 & f_1 & \dots & f_P \\ f_1 & f_2 & \dots & f_{P+1} \\ \vdots & \vdots & \vdots & \vdots \\ f_{D-P-1} & f_{D-P} & \dots & f_{D-1} \end{bmatrix} \begin{bmatrix} \hat{a}_0 \\ \hat{a}_1 \\ \vdots \\ \hat{a}_P \end{bmatrix} = 0$$

(3.4a)

where the vector of unknowns is the set of estimated coefficients of the polynomial in Eq. (3.3a) (sometimes called the characteristic equation). Observe that Eq. (3.4a) has the form of a linear predictor or difference equation, which occurs commonly in various signal processing procedures.

We see that a constraint or auxiliary equation is needed to make this system of equations inhomogeneous. It is convenient to impose the condition that the p'th coefficient of the polynomial in Eq. (3.3a) is unity, i.e.

$$\sum_{\alpha=1}^{P-1} \hat{a}_{\alpha} Z^{\alpha} = -Z^P$$

which can then be written in terms of the roots Z_i as

(3.3b)

$$\sum_{\alpha=1}^P \hat{a}_{\alpha} Z^{\alpha} = (Z - \hat{Z}_1)(Z - \hat{Z}_2) \dots (Z - \hat{Z}_P) = \prod_{\alpha=1}^P (Z - \hat{Z}_{\alpha})$$

leading to

$$\begin{bmatrix} f_0 & f_1 & \dots & f_{P-1} \\ f_1 & f_2 & \dots & f_P \\ \vdots & \vdots & \vdots & \vdots \\ f_{D-P-1} & f_{D-P} & \dots & f_{D-2} \end{bmatrix} \begin{bmatrix} \hat{a}_0 \\ \hat{a}_1 \\ \vdots \\ \hat{a}_{P-1} \end{bmatrix} = - \begin{bmatrix} f_P \\ f_{P+1} \\ \vdots \\ f_{D-1} \end{bmatrix}$$

(3.4b)

Other constraints could also be used, e.g.,

$$\sum_{\alpha=0}^P \hat{a}_{\alpha}^2 = 1.$$

The linear system in Eq. (3.4b) provides the basis for obtaining the P coefficients of the polynomial in Eq. (3.3b), from whose roots the P pole values are found as

$$\hat{s}_{\alpha} = (1/\delta) \ln(\hat{Z}_{\alpha})$$

(3.5)

If $D = 2P$, the system is square of order P and can be solved directly for the P characteristic-equation coefficients. For $D > 2P$, the method of least squares can be used. In either case, the residues can be found by various means, such as fitting the first P equations of Eq. (3.2) or by applying the least-squares technique to the entire set. When the correct (or most appropriate) value for P is unknown, then it may be necessary to experiment using different values of the processing parameters P , D , and δ [2].

It is important to recognize that for the method to work the data must be available in uniform steps of the exponent, so that Eq. (2.1) can be written in polynomial form as in Eq. (3.1), which limits application of the this form of Prony's Method. Uniform sampling in the spectral variable is not a requirement of the spectral-domain version of Prony's Method, as will be shown below. Note also that the technique just outlined can be used for data whose exponential terms are products of two or more exponents depending on different variables, e.g., where

$$f(x,y) = \sum_{\alpha=1}^{W_x} \sum_{\beta=1}^{W_y} R_{\alpha\beta} \exp(s_{\alpha}^{(x)}) \exp(s_{\beta}^{(y)}) \quad (3.6)$$

a form that occurs in connection with rectangular-mesh sampling [2].

3.2.1 Some Results Obtained from Waveform-Function Sampling

—Time-series data

The original motivation for using Prony's Method in electromagnetic's applications was that of finding the complex-frequency resonances which arise in the Singularity-Expansion Method (SEM) [6]. In particular, since the poles and residues determine any late-time transient response it seemed reasonable to expect that these parameters should be accessible from various EM phenomena whether measured or computed. The feasibility for doing so in electromagnetics was first demonstrated by Van Blaricum and Mittra [12] whose success instigated much further work on this very basic problem (see for example Poggio et. al. [13], Van Blaricum and Mittra [14], and Miller [15]). For completeness in our present discussion, some examples of time-series processing are included as reported in [15].

Aside from the analytical utility of SEM poles, there is also the physical insight that can be derived therefrom, as demonstrated in Figure 3.1 [12]. Presented there are the poles estimated from waveform-domain parameter estimation applied to the computed transient response of a center-loaded straight wire, and an unloaded wire bent at its center at a right angle. We observe that for the former the $n = 1, 3, 5, \dots$ and for the latter the $n = 2, 4, 6, \dots$ poles are more lossy. The increase in the damping component of the pole is due to the increased resistive loss caused by the I^2R of the current maximum at the center load and to the increased radiative loss caused by acceleration-induced radiation of the charge maximum at the center bend, respectively. While an explanation for such losses might be deduced in other ways, the pole behavior provides a physically intuitive demonstration of basic electromagnetic phenomenology.

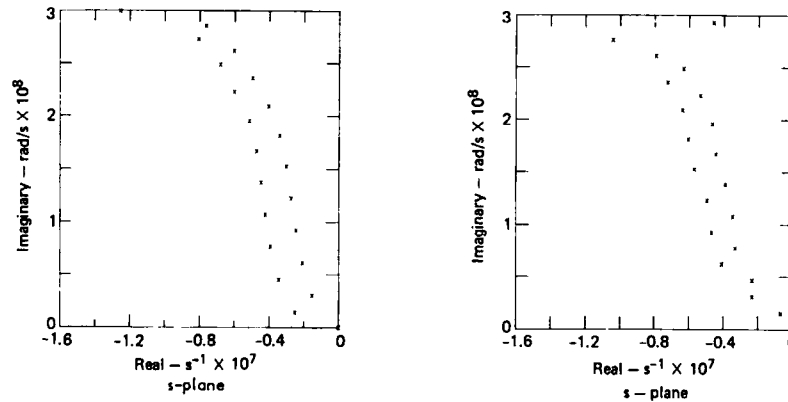


Figure 3.1. Poles estimated from computed time responses for a center-loaded straight wire, part (a) and a wire of equal length with a 90-degree bend at its center, part (b). The increased damping components of poles 1, 3, ... for the former and of poles 2, 4, ... for the latter are due to the dissipative and radiative loss of the corresponding current/charge modes on these objects. It can be concluded that the poles represent fundamental physical phenomena which characterize an object's geometrical and electrical properties, with the oscillatory component being more sensitive to size and the damping component being more sensitive to shape and impedance.

Another example of waveform-domain parameter estimation is given in Figure 3.2. Presented there are the superimposed waveforms of the fields scattered from a straight wire and circular loop, both having the same overall wire length, when excited by a Gaussian-pulse plane wave together with the corresponding estimated poles. We observe that the loop's response contains only half as many poles as the straight wire due to the fact that the former can support only modes having integer numbers of wavelengths whereas the latter supports both half-integer and integer modes. It can also be seen that although the oscillatory components of the integer-wavelength resonances are approximately equal, the damping components for the loop are more than twice as large. This is due to the fact that charges on the straight wire primarily experience acceleration only at its ends while those on the loop undergo continuous acceleration as they move around a continuously curved path. We can observe in general that the oscillatory pole component is a more sensitive measure of object size while the damping component conversely is a more sensitive measure of object shape. When we observe that noise produces more uncertainty in the value of the damping component, it can be concluded that the use of poles for target identification would be more reliable for size discrimination than would be the case for shape discrimination [2].

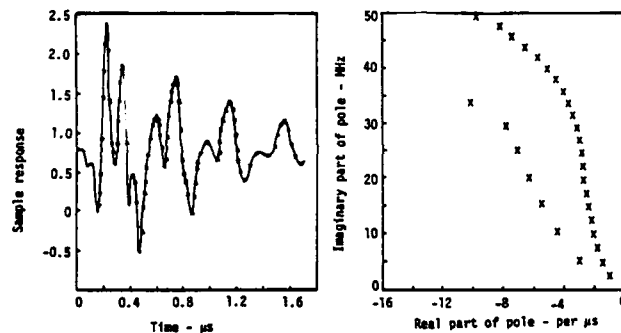


Figure 3.2. Computed time waveform and the estimated poles for Gaussian-pulse scattering from electrically noninteracting straight wire and circular loop of the same overall length. The increased radiative loss due the loop's curvature is clearly evident as well as the fact that since it supports only the odd current modes, it has only half the poles per unit bandwidth.

—Far-field patterns

The far fields produced by various source geometries are also candidates for waveform-domain parameter estimation. One example is illustrated in Figure 3.3 where the farfield (radiation pattern) of a linear array of 11 point sources is analyzed for aperture widths varying from one to nine wavelengths [15]. As the aperture width increases, it can be seen that the estimation process produces more accurate values for the source locations, stabilizing after the source separation exceeds one-half wavelength. An explanation for this result is demonstrated in Figure 3.4, where the eigenvalue spectrum of the data matrix [see Eq. (3.4b) above] is seen to become broader with increasing aperture width, or equivalently that the condition number is becoming smaller. This means essentially that the source locations are becoming more "observable" and thereby more accurately computable. Because there are only 11 sources contributing to the pattern being analyzed, the 12th eigenvalue would remain small whatever aperture width is used, showing that the eigenvalue spectrum provides a measure of the pattern rank. Eigenvalue spectra like this are typical of waveform-domain parameter estimation, and demonstrate the relationship between the accuracy of the parameter estimates and the linear independency (or lack thereof) of the data equations from which they are obtained.

An example of using waveform-domain parameter estimation for pattern synthesis is presented in Figure 3.5. The prescribed pattern magnitude is shown together with the source distribution derived for $P = 13$, the pattern produced by that distribution, and the "constrained" pattern which results from setting the real components, $s_{\omega r}$, of the source locations to zero. Nonzero values for $s_{\omega r}$ in our notation correspond to imaginary locations in real space and give rise to a source directivity proportional to $\exp[s_{\omega r} \cos(\theta)]$. Thus, by setting $s_{\omega r}$ to zero the effect on synthesizing the desired pattern by using only isotropic sources, as implicitly assumed in our approach, can be tested. The difference between the constrained and unconstrained patterns is seen to be small in this case, showing the pattern to be effectively realizable using MBPE.

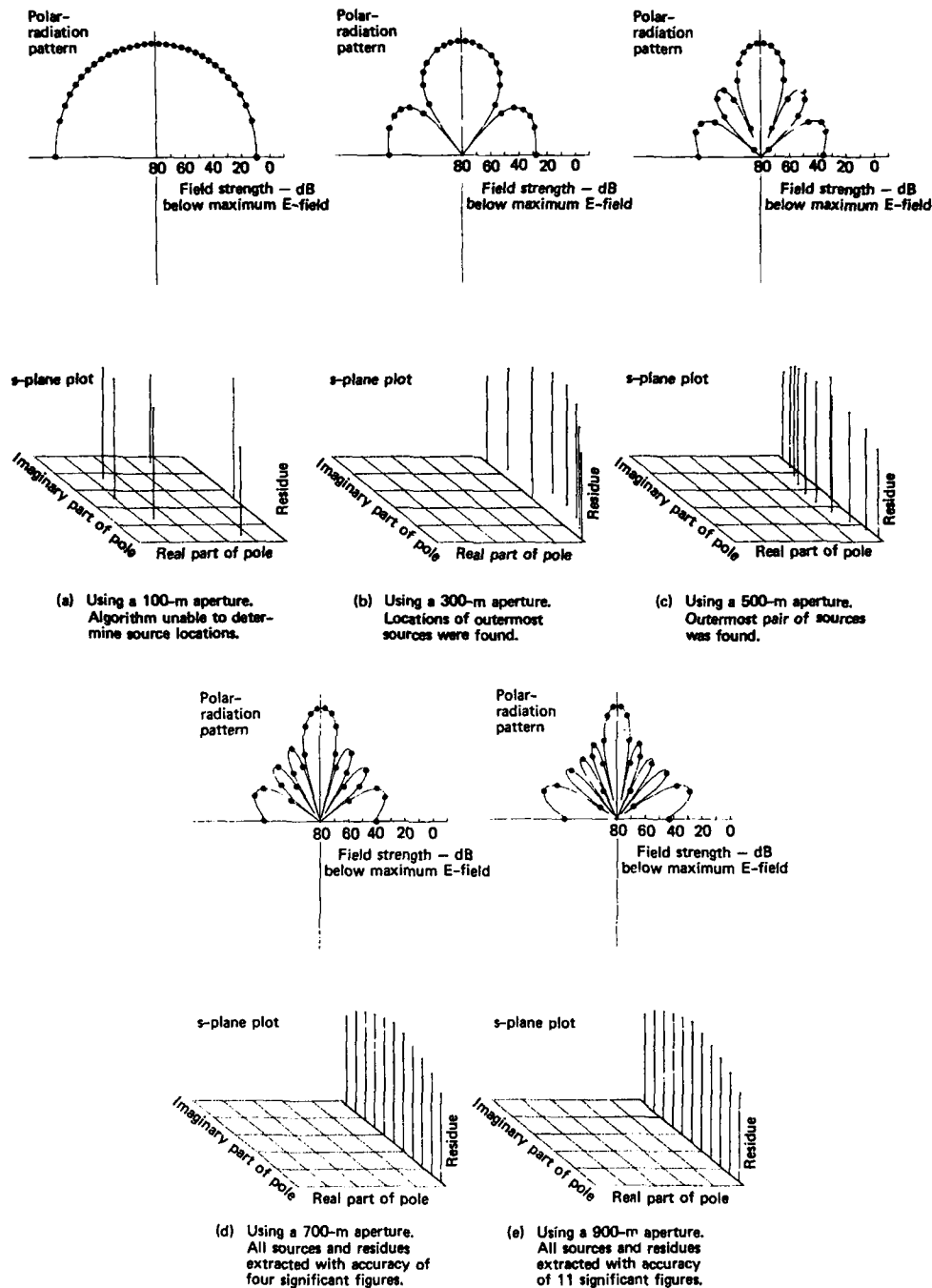


Figure 3.3. The radiation patterns [given by $\sum \exp[ik d_n \cos(\theta)]$] and estimated poles for an 11-source array of point radiators as a function of aperture width for a wavelength of 100 m. The source locations d_n in complex space are related to the pole value s_n by $d_n = s_n / ik$ where k is the wavenumber. The real-pole component should be zero, but for smaller aperture widths is not because the sources are not all resolvable, thus implying a source directivity not included in the original pattern calculation. For this case, the sampling is done in uniform steps of $\cos(\theta)$ as required for the polynomial form of Equation (3.1) to be obtained.

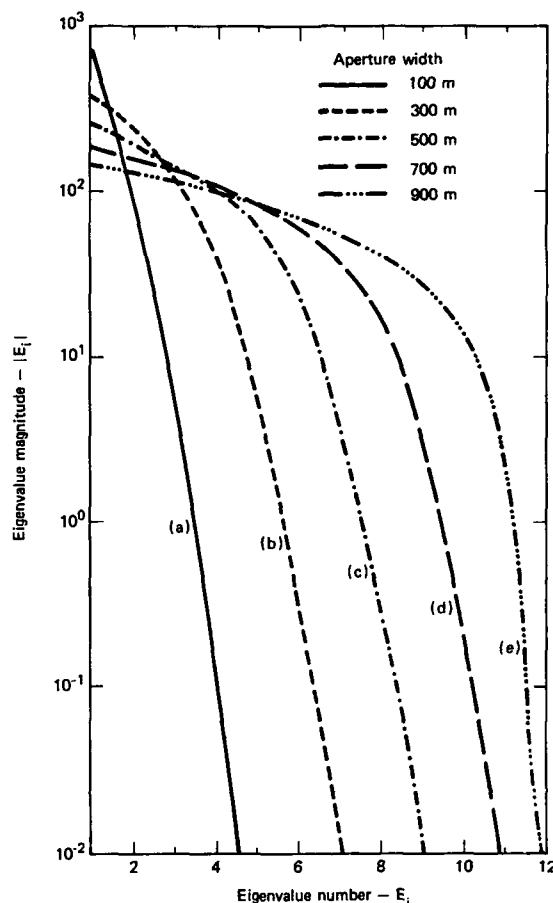


Figure 3.4. Eigenvalue spectra for the problem of Figure 3.3. The increasing observability of the source poles is indicated by the spreading out of the spectra with increasing aperture width and the corresponding decrease in the spectrum slope out to the 11'th eigen or wavenumber. Although the nominal condition number of the data matrix remains rather large even for the 900 m aperture, its effective value, in terms of the ratio of EV_1/EV_{11} , is quite reasonable.

3.3 Using Waveform-Derivative Sampling

If the waveform from which the parameters are to be estimated is sampled in terms of its derivatives rather than as a sequence of equally-space observations, the basic procedure just outlined can also be employed. While the use of waveform-derivative sampling might seem unusual, it is included for completeness and because spectral-derivative sampling offers some attractive advantages [7] which might also eventually be found for the former as well. We begin by writing the i 'th derivative of Eq. (2.1) without the nonpole term as

$$f^{(i)} = \frac{d^i f(x)}{dx^i} = \sum_{\alpha=1}^W s_{\alpha}^i R_{\alpha} \exp(s_{\alpha} x) \quad (3.7)$$

Proceeding in the same way as for the case of waveform-function sampling, but using the polynomial

$$\sum_{\alpha=1}^P b_{\alpha} s^{\alpha} = 0; s = s_i, i = 1, 2, \dots, P$$

we can derive the following linear system for the predictor coefficients

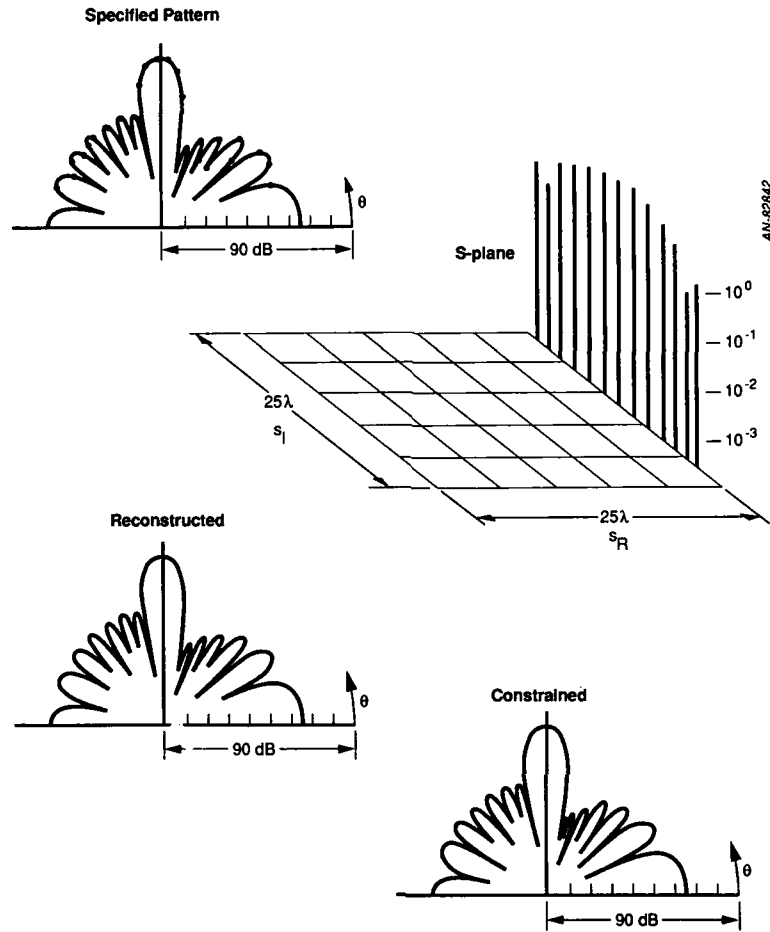


Figure 3.5. Example of using waveform-domain MBPE for pattern synthesis. The desired pattern is shown in (a), the source distribution using $P = 13$ is shown in a perspective plot in (b), and the synthesized pattern and constrained pattern (obtained by setting $s_m = 0$) in (c) and (d). Note that only the number of sources to be used in the synthesis is specified *a priori*, with both the source positions and strengths derived from MBPE. This leads in general, to a nonuniform spacing of the discrete sources used to synthesize a continuous aperture distribution.

$$\begin{bmatrix} f^{(0)} & f^{(1)} & \dots & f^{(P-1)} \\ f^{(1)} & f^{(2)} & \dots & f^{(P)} \\ \vdots & \vdots & \ddots & \vdots \\ f^{(D-1)} & f^{(D-P)} & \dots & f^{(D-2)} \end{bmatrix} \begin{bmatrix} \hat{b}_0 \\ \hat{b}_1 \\ \vdots \\ \hat{b}_{P-1} \end{bmatrix} = - \begin{bmatrix} f^{(P)} \\ f^{(P+1)} \\ \vdots \\ f^{(D-1)} \end{bmatrix} \quad (3.8)$$

which is the counterpart of Eq. (3.4b) with function samples replaced by derivative samples and where we have also set the P 'th coefficient in the characteristic equation to unity. But where the poles are found from logarithmic transforms of the roots of the characteristic equation for waveform-function sampling, we now find that they are themselves the roots of the characteristic equation for waveform-derivative sampling.

3.3.1 Some Results Obtained from Waveform-Derivative Sampling

A sample result for waveform-derivative sampling of an 8-pole waveform is given in Table II below. These results are similar to those obtained from waveform-function sampling in that the real, or damping component of the pole is less accurately obtained than is the oscillatory or imaginary component.

TABLE 2
TEST RESULTS FOR WAVEFORM-DERIVATIVE SAMPLING

INPUT POLES		ESTIMATED POLES	
Real	Imaginary	Real	Imaginary
-1.000	-.3536	-.0403	-.3660
-1.000	.3536	-.0403	.3660
-1.000	-1.0607	-.0735	-1.0617
-1.000	1.0607	-.0735	1.0617
-1.000	-1.7678	-.1062	-1.7680
-1.000	1.7678	-.1062	1.7680
-1.000	-2.4749	-.0978	-2.4753
-1.000	2.4749	-.0978	2.4753

4 PROCESSING IN THE SPECTRAL DOMAIN

4.1 Introduction

Waveform-domain sampling starts with an exponential-series generating function whereas spectral-domain sampling starts with its pole-series counterpart. While this might seem to be a minor difference, it is significant in that the need for uniformly-spaced samples that applies in waveform sampling is not a requirement in the spectral domain. This is because the spectral-domain model is already of polynomial form as will be seen below, a potential advantage. On the other hand, the spectral-domain linear system whose solution yields the model parameters has its data samples multiplied by the spectral variable raised to integer powers from 0 to P , producing a possibly very large dynamic range in the matrix coefficients. As one potential compensation for the latter possibility, it can be feasible to reduce the model order, i.e., the number P of poles in the estimator, below the actual number of poles W in the generator by using data from a reduced spectral range in the estimation process. We discuss spectral-function sampling and spectral-derivative sampling in turn below.

4.2 Using Spectral-Function Sampling

Determination from spectral-function sampling of the poles in Eq. (2.2) using the pole-series version of Prony's method involves sampling $F(X)$ at D data points $X = X_i$, $i = 0, 1, \dots, D-1$ over a spectral interval of width $\Delta = \text{Im}(X_{D-1} - X_0)$, from which a linear system of equations is also derived. But whereas in the waveform (time) domain only the free-response portion of the waveform can be modeled because the function samples used are restricted to the time after the excitation is no longer affecting the response, this is not strictly feasible in the spectral domain. Although reasonable results can be found using only the pole-based part of Eq. (2.2) [8], more robust performance will be obtained if the nonpole portion is modeled as well, which is analogous to including the driven part of the response in the waveform domain. Thus, with $X = \sigma + j\omega$ a general complex frequency (where sampling on the $j\omega$ axis would require $\sigma = 0$) we have

$$F(X) = F_p(X) + F_{np}(X) = \sum_{\alpha=1}^W \frac{R_\alpha}{(X - s_\alpha)} + \sum_{\beta=-Q}^R C_\beta X^\beta \quad (4.1)$$

Note that the nonpole part of Eq. (4.1) is written to have polynomial terms in complex frequency extending from powers of $-Q$ to $+R$ and that for most practical cases, the value of Q would be limited to 0 or 1. Also observe that while the number of poles in an electromagnetic response function is theoretically infinite, a computationally impractical situation, here we are concerned with a finite spectral range so that the number of poles needed in the model can be relatively small assuming the effect of those poles outside that range is either negligible or can be appropriately taken into account. This also helps avoid problems associated with ill conditioning that might arise when too many poles and too wide a spectral range are covered, but does presume that two or more different models might then be used to cover a wider range of interest. Use of two (or more) models yields a second benefit, as the difference between models in overlapping frequency intervals could then provide an error measure for an adaptive-sampling scheme. Finally, it should be observed that the nonpole term in Eq. (4.1) can also help account for those poles outside the observation interval which affect the spectral samples taken within it, since for poles such that $s_\alpha \approx X$, these terms behave essentially as $1/X$ while for $s_\alpha \approx X$, these terms are essentially constants proportional to $1/s_\alpha$.

Upon transforming Eq. (4.1) into least-common denominator form, we obtain the rational-function counterpart

$$F(X) = N(X)/D(X) \quad (4.2a)$$

for which a more convenient computational version is obtained by multiplying by $D(X)$ to get

$$F(X)D(X) = N(X) \quad (4.2b)$$

The numerator and denominator polynomials are given in turn by

$$N(X) = N_0 + N_1 X^1 + N_2 X^2 + \dots + N_n X^n \quad (4.3a)$$

and

$$D(X) = [D_0 + D_1 X^1 + D_2 X^2 + \dots + D_Q X^Q] X^0 \quad (4.3b)$$

with terms of maximum order $n = W+Q+R$ and $d = W$ respectively.

Although there seem to be $n+d+2 = 2W+Q+R+2$ unknown coefficients in Eq. (4.3) as written, as before at least one of these must be specified independently to obtain an inhomogeneous set of equations, a conclusion that can also be reached by noting that there are only $2W+Q+R+1$ unknown coefficients represented by the $W R_n$ and s_n and $(Q+R+1)C_p$. Furthermore, the $2W$ complex parameters represented by the R_n and s_n account for only $2W$ real unknowns because the poles and residues occur in conjugate pairs, which also means that the numerator- and denominator-polynomial coefficients are real as well. Although any one of the numerator- and denominator-polynomial coefficients could be specified, for spectral-function sampling we will set D_0 to unity analogous to what was already done for waveform sampling. We note that for the special cases that $F_{np}(X) = 0$ or $F_{np}(X) = \text{constant}$, then $Q = R = C_0 = 0$ and $Q = R = 0$, respectively. Because W may not be known *a priori*, its value might also need to be established for numerical purposes by conducting computer experiments as was already observed in the case of waveform-domain sampling.

The resulting procedure may be regarded as a logical extension of Prony's Method which was developed originally for an exponential, rather than a pole, series. For this reason the technique summarized below has been referred to as "Frequency-Domain Prony" [8], although it can also be described as rational-function approximation.

In sampled form, Eq. (4.2b) is given by

$$F_i \hat{D}(X_i) = \hat{N}(X_i)$$

which reduces to the following linear system [Miller (1987)]

$$\begin{bmatrix} F_0 & X_0 F_0 & \dots & X_0^{d-1} F_0 & -1 & -X_0 & \dots & -X_0^n \\ F_1 & X_1 F_1 & \dots & X_1^{d-1} F_1 & -1 & -X_1 & \dots & -X_1^n \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ F_{D-1} & X_{D-1} F_{D-1} & \dots & X_{D-1}^{d-1} F_{D-1} & -1 & -X_{D-1} & \dots & -X_{D-1}^n \end{bmatrix} \begin{bmatrix} \hat{D}_0 \\ \vdots \\ \hat{D}_{d-1} \\ \hat{N}_0 \\ \vdots \\ \hat{N}_n \end{bmatrix} = \begin{bmatrix} X_0^d F_0 \\ X_1^d F_1 \\ \vdots \\ \vdots \\ X_{D-2}^d F_{D-2} \\ X_{D-1}^d F_{D-1} \end{bmatrix} \quad (4.4)$$

where again the carat denotes estimated values and $D \geq 2P+Q+R+1$ is required.

The rational-function approximation to the sampled transfer function is thereby obtained in terms of the frequency-sampled data F_i by solving Eq. (4.4) for the numerator- and denominator-polynomial coefficients. It is interesting to observe that in contrast with the waveform-function and waveform-derivative models discussed above where the pole-related (denominator) coefficients are found independently from the residue solution, the spectral-function solution involves finding both together. At this point, the model calculation has been completed, although as a further step we could also obtain the system poles and zeroes by solving $D(s) = 0 = D(s_p)$ and $N(z) = 0 = N(z_p)$.

4.2.1 Some Results Obtained from Spectral-Function Sampling

—Frequency-domain transfer functions

Upon applying a procedure like that just described but using magnitude-only data [2] for the feed-point current and broadside radiated field for a center-excited straight-wire antenna, results of the kind presented in Figure 4.1 are obtained. Results for both positive and negative frequencies were employed as input data, from which, it can be observed, the pole locations are accurately obtained. This result is partly due to the fact that the structure being analyzed is highly resonant so that the poles are close to the imaginary axis. The computation performed here included all data points shown and so encompassed a wide frequency range, with the possibility of leading to a poorly-conditioned data matrix. It would generally be preferable to use lower-order models to circumvent this possibility (as discussed below), with a wider frequency range being covered by using multiple models and frequency "windows".

Another example of using frequency-domain data for MBPE is illustrated by the results presented in Figure 4.2 [2]. The frequency-dependent reflection coefficient for a normally incident plane wave on a two-layer dielectric slab is shown together with the poles which are estimated. In this case, each pole represents a wave reflected at one of the planar interfaces, with the oscillatory component increasing with the total electrical distance traversed by the given component. The poles have no damping component because the slabs are lossless, and their residues decrease because the wave amplitudes decrease with each successive bounce.

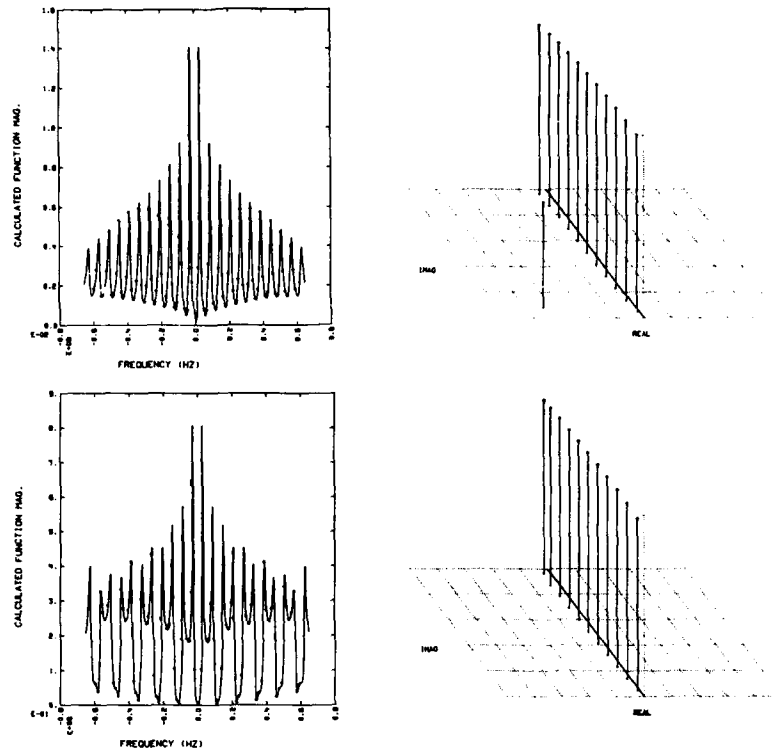


Figure 4.1. Results for a center-excited antenna. Shown in part (a) are the spectrum and estimated poles using the feedpoint current, and in part (b) the corresponding results for the broadside, far-field response. The dots indicate the sample points and the continuous curve the reconstructed transfer function.

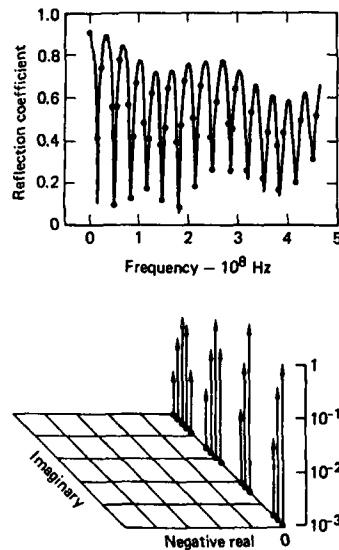


Figure 4.2. Results for a two-layer planar slab. The reflection coefficient for normal incidence is shown as part (a) and the estimated poles in part (b). The two-layer slab is embedded in free space, with the layer permittivities ϵ_1 and ϵ_2 relative to free space equal to 9 and 16 and with thickness given by D_1 and D_2 of 0.208 m and 1.00 m respectively.

As two last examples of MBPE applied to frequency-domain data, we present in Figures 4.3 and 4.4 results which demonstrate the use of multiple data sets for helping to separate curve-fitting poles from poles actually present in the data. Curve-fitting poles of course arise whenever the number of model poles exceeds the number actually present in the data. Various schemes might be somewhat useful in identifying to which set a given pole belongs. For example, poles in the right-half plane and poles which have very large damping components might both be rejected on physical grounds, while poles having very small residues might be discarded because they contribute insignificantly to a $j\omega$ -axis transfer function. Another approach which helps accomplish this separation but which at the same can provide an average value for the pole and estimate of its uncertainty is the use of multiple data sets. These can arise from employing subinterval frequency windows to scan over a wider frequency range, or from systematically varying the excitation as happens when the incidence angle of a plane wave is changed in steps over some range of values.

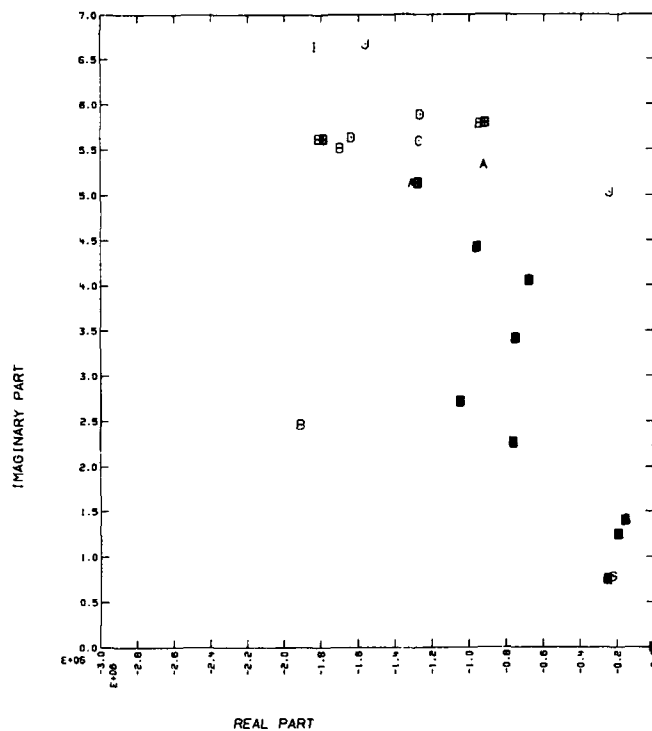


Figure 4.3. Computing object poles from multiple frequency windows. Results are shown for plane-wave incidence on a generic target whose backscattered field is modeled using 40 poles (20 complex-conjugate pairs) over a succession of 5-MHz-wide frequency windows. Results from individual windows are designated by A-J respectively, with the darker clusters indicating good agreement being found among the separate estimates.

The poles obtained from "sliding" a 5-MHz window along the $j\omega$ axis to obtain 10 separate sets of poles for a generic target excited by an incident plane wave, using a total of 40 poles (20 pairs) in the estimation model, are shown in the second quadrant of the complex-frequency plane in Figure 4.3. All the poles computed are plotted, designated by the letters A to J, to identify those poles obtained from a given window. The darker symbols are overlays of several, in some case as many as all 10, computed poles indicating the numerical agreement that these separate calculations can yield. Results for 10 incidence angles from 25° to 65° over a frequency range of 0-5 MHz are presented in a similar fashion in Figure 4.4. Again, tightly clustered sets of poles are obtained, where furthermore good agreement is found with the sliding-window results over their mutual frequency range of 0-5 MHz. These results demonstrate the benefit of appropriately using redundant data to help identify the actual EM poles when that is the goal of the modeling exercise.

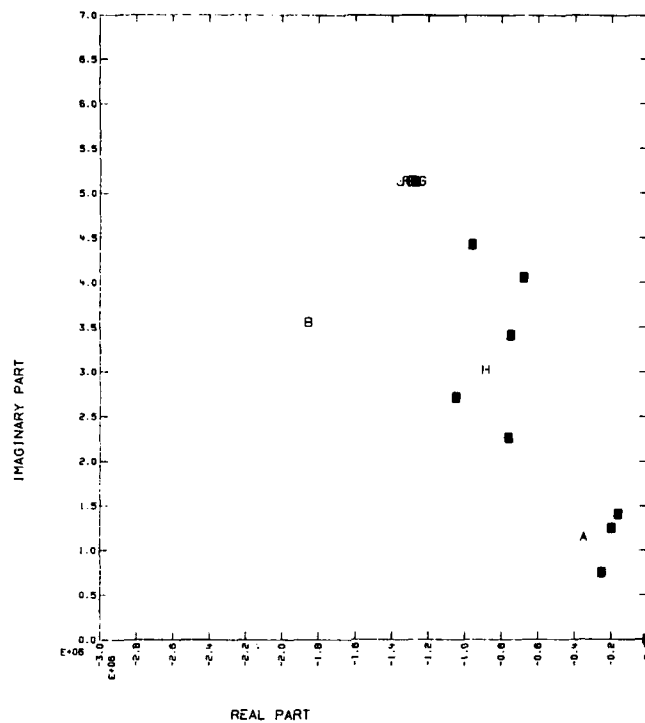


Figure 4.4. Computing object poles from multiple excitations. A single 0-5 MHz window is used for 10 angles of plane-wave incidence between 25° and 65°, again using letters A-J to designate the individual results. The poles found from different excitations exhibit agreement comparable to that found for different frequency windows in Figure 4.4, with good agreement also found between the results of these two multiple-processing exercises.

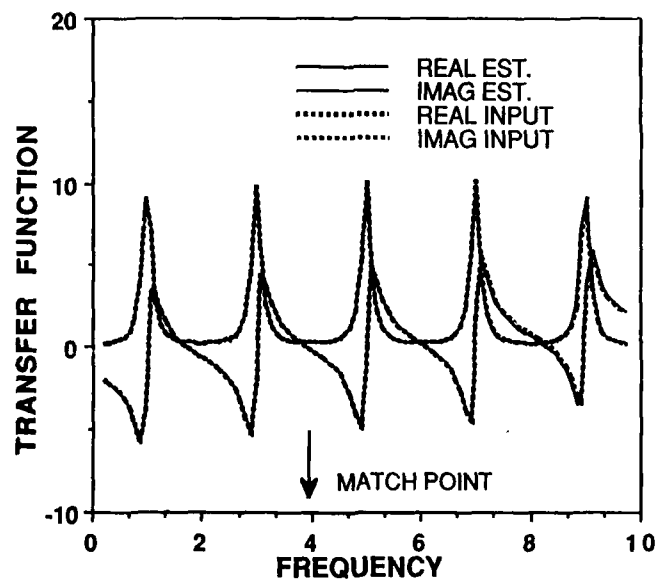


Figure 4.5. Results for a simulated 5-pole spectrum. The input data was the response and its first nine derivatives at 4 on the frequency axis.

4.3 Using Spectral-Derivative Sampling

The same rational-function model can be used when frequency-derivative samples are employed rather than frequency samples. Starting with Eq. (4.2) and upon differentiating t times with respect to X and dropping the explicit X dependence, there results the following equations

$$\begin{aligned} FD &= N \\ F'D + FD' &= N' \\ F''D + 2F'D' + FD'' &= N'' \\ F'''D + 3F''D' + 3F'D'' + FD''' &= N''' \\ &\vdots \\ F^{(t)}D + tF^{(t-1)}D' + \dots + C_{t,m}F^{(m)}D^{(t-m)} + \dots + FD^{(t)} &= N^{(t)} \end{aligned} \quad (4.5)$$

where $C_{r,s} = r!/[s!(r-s)!]$ is the binomial coefficient. The system of $t+1$ equations in Eq. (4.5) (there are t frequency-derivative samples plus the original frequency sample) provides the information from which the model coefficients can be found if $t \geq n+d+1$, assuming as before that one of the coefficients is specified. With no loss of generality, Eq. (4.5) can be simplified by setting $X = 0$ since X can represent the variation in spectral variable about the point where the spectral-derivative samples are obtained. We then find the following matrix representation for the unknown numerator- and denominator-polynomial coefficients

$$\begin{bmatrix} 1 & 0 & \dots & 0 & 0 & \dots & 0 \\ 0 & 1 & \dots & -F_0 & 0 & \dots & 0 \\ 0 & 0 & \dots & -F_1 & -F_0 & \dots & 0 \\ 0 & 0 & \dots & 1 & -F_2 & -F_1 & \vdots \\ 0 & 0 & \dots & 0 & \vdots & -F_2 & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & 0 \\ 0 & 0 & \dots & 0 & -F_{D-1} & -F_{D-2} & \dots \end{bmatrix} \begin{bmatrix} \hat{N}_0 \\ \hat{N}_1 \\ \vdots \\ \hat{N}_n \\ \hat{D}_1 \\ \vdots \\ \hat{D}_d \end{bmatrix} = \begin{bmatrix} F_0 \\ F_1 \\ F_2 \\ \vdots \\ \vdots \\ \vdots \\ F_{D-1} \end{bmatrix} \quad (4.6)$$

where the samples are now $F_m = F^{(m)}/m!$. Again we note that at least one coefficient of the two polynomials must be specified to render the system inhomogeneous. We have chosen here to use $D_0 = 1$. As was the case for spectral-function sampling, spectral-derivative sampling also produces a linear system which contains both the numerator and denominator coefficients.

If the sampling is performed at two (or more) frequencies, then the expansion about $X = 0$ cannot be used, and the system matrix takes the more general form

$$[D]\hat{X} = F$$

where

$$[D] = \begin{bmatrix} 1 & X_1 & X_1^2 & \dots & X_1^n & -X_1 F_{0,1} & -X_1^2 F_{0,1} & \dots & -X_1^d F_{0,1} \\ 0 & 1 & 2X_1 & \dots & nX_1^{n-1} & -\sum_{i=0}^1 X_1^i F_{1-1} M_{1,1} & -\sum_{i=0}^2 X_1^i F_{1-1} M_{1,2} & \dots & -\sum_{i=0}^d X_1^i F_{1-1} M_{1,d} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & \frac{n!}{(n-1)!} X_1^{n-1} & -\sum_{i=0}^1 X_1^i F_{1-1} M_{1,1} & -\sum_{i=0}^2 X_1^i F_{1-1} M_{1,2} & \dots & -\sum_{i=0}^{t,d} X_1^i F_{1-1} M_{1,2} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & X_p & X_p^2 & \dots & X_p^n & -X_p F_{0,p} & -X_p^2 F_{0,p} & \dots & -X_p^d F_{0,p} \\ 0 & 1 & 2X_p & \dots & nX_p^{n-1} & -\sum_{i=0}^1 X_p^i F_{1-p} M_{1,1} & -\sum_{i=0}^2 X_p^i F_{1-p} M_{1,2} & \dots & -\sum_{i=0}^d X_p^i F_{1-p} M_{1,d} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & \frac{n!}{(n-1)!} X_p^{n-1} & -\sum_{i=0}^1 X_p^i F_{1-p} M_{1,1} & -\sum_{i=0}^2 X_p^i F_{1-p} M_{1,2} & \dots & -\sum_{i=0}^{t,d} X_p^i F_{1-p} M_{1,d} \end{bmatrix} \quad (4.7a)$$

and

$$\hat{X} = \begin{bmatrix} \hat{N}_0 \\ \hat{N}_1 \\ \vdots \\ \hat{N}_n \\ \hat{D}_1 \\ \hat{D}_2 \\ \vdots \\ \hat{D}_d \end{bmatrix}, \text{ and } F = \begin{bmatrix} F_{0,1} \\ F_{1,1} \\ \vdots \\ F_{t,1} \\ \vdots \\ F_{0,p} \\ F_{1,p} \\ \vdots \\ F_{t,p} \end{bmatrix} \quad (4.7b,c)$$

with

$$M_{i,j,k} = i!k!j!/(i-j)!/(k-j)!$$

In the above equations, the first subscript on the data sample F_{ij} denotes the order of the frequency derivative up to the maximum of t_i and the second denotes the frequency j at which the sample is obtained. The upper limit on the summations is the lower of t_i or d .

Observe that the elements in the matrix resulting from spectral-derivative sampling might involve a smaller dynamic range because the spectral variable is not raised to a maximum power of P as is the case of spectral-function sampling. This difference might appear to produce a better-conditioned matrix. However, besides the $m!$ multiplier of the m 'th spectral derivative, in the vicinity of a pole the derivatives themselves can become proportional to a large number raised to the $(n+1)$ st power, as differentiating the pole part of Eq. (2.1) reveals, i.e.,

$$\begin{aligned} F^{(n)} &= d^n [\Sigma R_\alpha / (s_\alpha - j\omega)] / d\omega^n \\ &= j^n / (-j\omega)^{n+1} \Sigma R_\alpha + R_j^n / (s_j - j\omega)^{n+1} + R_j^n / (s_j - j\omega)^{n+1} \\ &\quad + j^n \Sigma R_\alpha / (s_\alpha)^{n+1}; \alpha = 1, \dots, P; s_{j,m} \leq j\omega \leq s_{j,m} \end{aligned} \quad (4.8)$$

Thus, the maximum number of poles that can be employed in the spectral estimator model might be limited by the dynamic range of the coefficients which appear in the model matrix for both function and derivative sampling. In either case, however, a compensating factor is the possibility that a lower-order (i.e., reduced number of model poles) model can be used because the data samples are most affected by the nearest poles of the spectrum being sampled.

This approach based on spectral-derivative sampling is similar to rational-function, or Pade, approximation [9], although we might regard it as a further generalization of the various extensions of Prony's Method as discussed above and elsewhere [2]. All of these variations of the basic Prony Method have the goal of using computed or measured data to develop a numerical model for some phenomenological response which can be approximated by either a series of exponentials or a series of poles, although as noted above in discussing Eq. (4.1), some nonpole (or nonexponential) terms can also be included as well, where $R \neq 0$ would be used. In the next section, we discuss how the frequency derivatives of the response can be found as a generalization of a moment-method model.

4.3.1 An Example of Spectral-Derivative Sampling Using a Simulated Transfer Function

Application of spectral- or frequency-derivative sampling to simulated electromagnetics data is illustrated in Figure 4.5. Presented there are the real and imaginary components of the input spectrum (dotted curves), and the corresponding estimated spectrum (solid lines) as obtained from using a single sample of the input spectrum and its first nine derivatives at a frequency of 4 (the units are arbitrary here). In this case, the rational-function model employed a numerator order $n = 4$ and a denominator $d = 5$ for a total of 10 coefficients needed for the estimation model (recall that the first term of the denominator polynomial is set equal to one). Results estimated for the poles and zeroes are in good (3-place or more) agreement with the input data, as are the spectra themselves which are graphically indistinguishable. Presentation of results for actual EM data are deferred until the following section where we discuss computing frequency derivatives in a moment-method model.

5 COMPUTING FREQUENCY DERIVATIVES IN A MOMENT-METHOD MODEL

5.1 The General Idea

As noted above, the parameters needed to numerically quantify the waveform-domain or spectral-domain models can in principle be obtained from whatever information is available. Because most moment-method modeling in electromagnetics based on integral equations is done in the frequency domain, we will limit subsequent discussion to spectral- (or frequency-) domain applications of MBPE. As an alternative to using only frequency sampling, which although straightforward offers no further computational efficiency beyond reducing the number of actual samples required to reconstruct a transfer function, we demonstrate the advantage of using frequency derivatives of the response. The basis for the technique is due to earlier work by Burke as reported in [10]. It can be shown [7] that the additional programming complexity that arises in computing frequency derivatives from a moment-method model can be compensated by a more than proportional reduction in overall computer time.

On writing the moment-method equations that arise from an integral-equation formulation in matrix form, we have the impedance equation

$$\sum_{j=1}^N Z_{ij}(\omega) I_j(\omega) = V_i(\omega) \quad (5.1a)$$

where $Z(\omega)$, $I(\omega)$ and $V(\omega)$ are the usual frequency-dependent impedance matrix and the current and voltage vectors, respectively. A solution for the current can then be formally written as an admittance relation

$$I_i(\omega) = \sum_{j=1}^N Y_{ij}(\omega) V_j(\omega) \quad (5.1b)$$

where the matrix $Y(\omega)$ is the inverse (or equivalent) of $Z(\omega)$. We note that the MBPE approach discussed here, whether using frequency-function or frequency-derivative sampling, could be implemented using LU factorization, iteration, or any other method to obtain a numerical solution of the linear system in Eq. (5.1a).

Because a solution for $I(\omega)$ or the fields it produces may be needed over a frequency range which spans many poles or resonances of $Y(\omega)$, it is usual to repeat these steps at enough frequency samples that a smooth representation of the transfer function can be derived. We note that each computation essentially repeats the entire solution process, resulting in a computer time generally proportional to $AN^2 + BN^3$ the terms of which account for filling the $Z(\omega)$ matrix and solving the $Y(\omega)$ respectively. The overall computer time is therefore approximated by $F(AN^2 + BN^3)$ if F frequency samples are needed and a frequency-independent representation of the object being modeled is used over the entire band. It is therefore advantageous not only to minimize the number of unknowns N consistent with accuracy requirements, but also to minimize F as well.

Two possibilities for working towards the goal of reducing, if not absolutely minimizing, F are discussed below. As might be expected from the preceding presentation, one approach is to employ the spectral-function model and sampling, and the other is to use a spectral-derivative model and sampling. We emphasize again that the analytical model employed for estimating the transfer function can be derived from whatever kinds of samples, function or derivative, are available.

Now let us differentiate the impedance equation with respect to frequency from which we obtain

$$\sum_{j=1}^N [Z_{ij}'(\omega) I_j(\omega) + Z_{ij}(\omega) I_j'(\omega)] = V_i'(\omega) \quad (5.2a)$$

where the prime denotes a frequency derivative. A solution of the differentiated impedance equation for the differentiated current can then be written

$$I_i'(\omega) = \sum_{j=1}^N Y_{ij}(\omega) \left(V_j'(\omega) - \sum_{k=1}^N Z_{ik}'(\omega) I_k(\omega) \right) \quad (5.2b)$$

where we observe that while the differentiated impedance matrix appears as part of a modified right hand-side of the differentiated admittance equation, $I'(\omega)$ is given in terms of an undifferentiated admittance matrix. Computing the differentiated current thus requires an additional number of computations beyond those needed for solution of the undifferentiated current proportional to N^2 rather than the N^3 that would apply to obtain another frequency sample (assuming that LU decomposition is used rather than iteration).

Continuing this process, and omitting the explicit frequency dependences for clarity, we find that the n 'th frequency derivative of the current is given by

$$I_i^{(n)} = \sum_{j=1}^N Y_{i,j} \left[V_j^{(n)} - \sum_{m=1}^n C_{n,m} \left(\sum_{k=1}^N Z_{i,k}^{(m)} I_k^{(n-m)} \right) \right] \quad (5.3)$$

where again $C_{n,k}$ is the binomial coefficient and the superscript in parenthesis indicates differentiation with respect to frequency of the order indicated. This equation is essentially identical in form to Eq. (4.5), and emphasizes the similarity between the differentiated rational-function and the moment-method models.

It is especially important to observe that information about the n 'th frequency derivative of the current continues to require a computer time proportional to N^2 . Expressed in another way, each additional frequency derivative of the solution vector for the current can be computed in a number of operations proportional to $A(n, N_{\text{max}})/N$ where A is a function which depends on the order of the derivative and the number of right-hand-sides for which the solution is sought. If the frequency derivatives provide information comparable to that available from the frequency samples themselves, it can be appreciated that there could be a substantial computational advantage to using the solution derivatives in estimating the transfer functions.

5.2 An Implementation Using NEC

To implement the procedure outlined above, which we denote as frequency-derivative interpolation (FDI), the derivatives of the impedance matrix must be evaluated. In a code such as NEC, the fields which comprise the impedance matrix can have two sources of frequency dependence. One of these is due to the frequency dependence of the field expressions themselves, and the other can arise, as in NEC, from the basis and testing functions. NEC employs point sampling, i.e., delta-function testing, which is not frequency dependent but does use the three-term sinusoidal current expansion function

$$I_i(s) = A_i + B_i \sin[k_s(s-s_i)] + C_i \cos[k_s(s-s_i)] \quad (5.4)$$

where k_s is normally set equal to the medium wavenumber k . This is done because analytical expressions are then available for five of the six field components needed for general problems. But the frequency dependence of the basis function which thereby occurs can greatly complicate implementation of FDI.

In implementing FDI in NEC it was decided to hold k_s constant in the basis functions to simplify the evaluation of derivatives. This step is justified by noting that the basis functions, although constructed from the functions of Eq. (5.4) which depend on k_s , have the form of B-splines which are insensitive to k_s for small $|k_s \delta|$. In addition, the solution is relatively insensitive to small changes in the shape of the basis functions. Hence, we started with equations for the field of the current components in Eq. (5.4) for $k_s \neq k$, derived the equations for the derivatives, and then set $k_s = k$. A further simplification was effected by using the VLF version of NEC (NEC3VLF) [11] in which the numerically canceling fields due to point charges at segment ends are analytically removed from the field equations. With the point-charge fields removed and $k_s \neq k$, the field components for a current of the form

$$I(z') = I_0 \begin{pmatrix} \sin(k_s z') \\ \cos(k_s z') \end{pmatrix} \quad (5.5a)$$

due to a segment extending from $-\delta$ to δ on the z axis of a cylindrical coordinate system are

$$E_p(\rho, z) = -\frac{j\eta I_0}{4\pi\rho} \left\{ \left[k_s(z-z') \begin{pmatrix} \cos k_s z' \\ -\sin k_s z' \end{pmatrix} \frac{e^{-jkR}}{kR} - j \begin{pmatrix} \sin k_s z' \\ \cos k_s z' \end{pmatrix} e^{-jkR} \right] \right\}_{-\delta}^{\delta} \\ - \frac{(k^2 - k_s^2)}{k} \int_{-\delta}^{\delta} dz' (z-z') \begin{pmatrix} \sin k_s z' \\ \cos k_s z' \end{pmatrix} \frac{e^{-jkR}}{R} \quad (5.5b)$$

and

$$E_z(\rho, z) = -\frac{j\eta I_0}{4\pi} \left\{ \left[-k_s \left(\frac{\cos k_s z'}{R} \right) \frac{e^{-jkR}}{kR} \right]_{-s}^s + \frac{(k^2 - k_s^2)}{k} \int_{-s}^s dz' \left(\frac{\sin k_s z'}{R} \right) \frac{e^{-jkR}}{R} \right\} \quad (5.5c)$$

The needed frequency derivatives of the fields are thus obtained by differentiating the above expressions with respect to k and then setting $k_s = k$. Consequently, derivatives of the integrals up to an order one less than the order of the derivative of the field remain when k_s is set equal to k . Considering in addition that E_z due to the constant term in the NEC basis is also needed, there are five integrals to evaluate for each order of differentiation.

The integrals and their derivatives were evaluated by expanding the exponentials and integrating the terms analytically. In this process each derivative shares some terms with the expansion for the previous derivatives. If numerical integration were used, the derivatives would share the $\exp(-jkR)$ term which accounts for much of the computation time. Hence, the additional effort is less than proportional to the number of derivatives. The remaining steps in evaluating the derivatives of Eq. (5.5) with respect to k are straightforward.

5.3 Some Results from NEC

The modified version of NEC described above was applied to three antenna problems as a preliminary test of the frequency-derivative approach. As one of the simpler examples that might be attempted, we chose a center-excited, straight-wire antenna. A more challenging case is presented by a straight horizontal wire located very near a conducting ground plane, which produces a much sharper resonance. The most difficult problem attempted is that of a forked monopole antenna having arms which are slightly different in length. Results from each case are summarized below.

5.3.1 Dipole Antenna in Free Space

Results are presented in Figure 5.1 for a center-excited dipole antenna modeled with NEC using 21 segments. Two rational-function models ($n = 4, d = 5$ on left and $n = 5, d = 4$ on right) were fit to the current and its first four derivatives at $L/\lambda_0 = 0.1$ and 2.0 and plotted using a dashed line. Directly computed values were obtained at 0.01 intervals in L/λ and are plotted for reference as a solid line. Both resonances are found accurately in the rational model, with that having the higher-order numerator producing the better result. We surmise that this is due to the fact that this model in effect includes not only a pole series but constant and linear terms in frequency. It is worth noting that using the same data in a Taylor series as that employed here for the rational-function model yields results whose accuracy monotonically decreases away from the match point as demonstrated in Figure 5.2. This emphasizes the importance in MBPE, when faced with choosing among alternate models, of selecting that which is most appropriate to the physics being modeled.

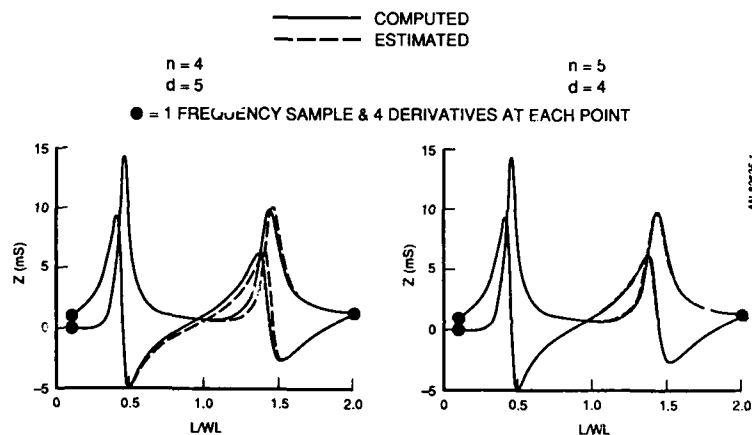


Figure 5.1. Input admittance for a dipole antenna of length L and thickness parameter $\Omega = 10$ in free space as computed from a moment-method model (—) and obtained using MBPE (---). Two results are shown, both based on samples of the input admittance and its first four derivatives at $L/\lambda = 0.1$ and 2.0 (indicated by the dots). In part (a) the rational-function model used $n = 4$ and $d = 5$ and in part (b) conversely $n = 5$ and $d = 4$. Although both models produce useful results, the latter is substantially better, apparently because the higher-order numerator polynomial includes a constant and linear frequency dependence in addition to simple poles.

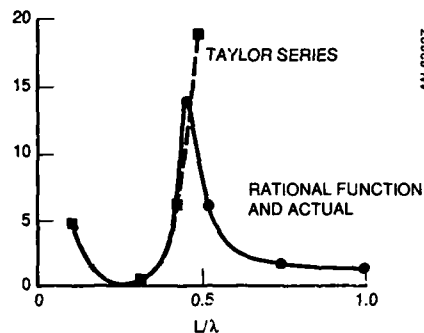


Figure 5.2. Taylor-series representation for a dipole antenna. The input conductance and its first four derivatives at $L/\lambda = 0.3$ is used in a Taylor-series expansion rather than the rational-function approximation, and demonstrates the limited radius of convergence due to the pole.

5.3.2 Horizontal Dipole Antenna Close to a Ground Plane

Results for a horizontal antenna near a perfectly-conducting ground plane are presented in Figure 5.3. This case was selected to test the frequency-derivative model for a response function having a higher-Q resonance. The solid curve was computed with a sampling interval of 0.001 in L/λ . It can be seen that for a relative bandwidth of about 10^{-3} , and using as model data the input admittance and its first four derivatives at $L/\lambda = 0.1$ (left plot, $n = d = 2$) and $L/\lambda = \pm 0.1$ (right plot, $n = 5, d = 4$), a good match with the actual results is obtained even though the sampling is done about 250 resonance widths away. We also observe that the MBPE resonance obtained using $n = d = 2$ is accurate to about 0.5 % in frequency, but when using the conjugate data as well, the match improves to within a few hundredths of a per cent.

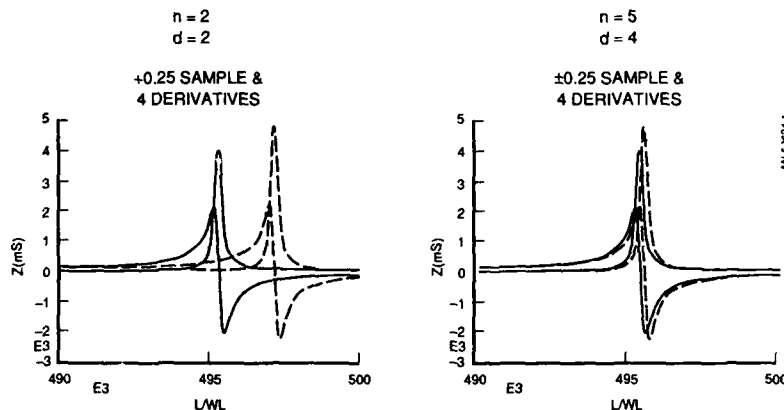


Figure 5.3. Dipole antenna of length L and thickness parameter $\Omega = 15$ at a height $L/50$ above a perfectly conducting ground plane. Directly computed values (—) are compared with the rational-function model (---) employing $n = 5$ and $d = 4$ based on the antenna admittance and its first four derivatives at $\pm j0.25$. The computed and model resonances agree to within a few hundredths of a per cent. Results obtained when using only the $+j0.25$ input data and $n = d = 2$ have about double this error in the resonance frequency.

5.3.3 Forked-Monopole Antenna

A still more challenging case is presented by a Vee or forked-monopole antenna having arms of unequal length where L is the length of the longer arm with the shorter arm of length $(14/15)L$ and with an included angle of 2.6 degrees, whose transfer function was presented previously in Figure 1.1. As repeated in Figure 5.4, this antenna is observed to have, in addition to a quarter-wave monopole resonance, a very sharp parasitic resonance due to a differential (transmission-line) current mode in the arms. The parasitic resonance corresponds to a pole very close to the real axis with a small residue, while the monopole resonance has a pole more distant from the real axis with a larger residue. This situation presents a severe challenge to the rational-function model which in this case employed the input admittance and its first four derivatives at $L/\lambda = 0.4$ and 0.8 and with $n = 5$ and $d = 4$. In spite of the extremely narrow parasitic resonance, the rational-function model matches the actual computations (sampled at 0.003 intervals in L/λ) so well that they are graphically indistinguishable.

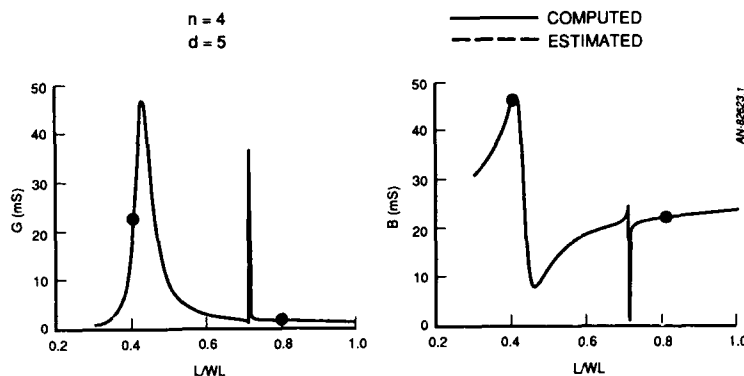


Figure 5.4. Input admittance of a forked monopole on a perfectly conducting ground. Directly computed values (_____) are compared with the rational-function model (_____) employing $n = 5$ and $d = 4$ based on the antenna admittance and its first four derivatives at $+j0.4$ and $+j0.8$ (indicated by the dots). In this case the computed and modeled results are graphically indistinguishable.

Results obtained from using only one derivative at each of 5 sampling frequencies, and when using 9 frequency samples and no derivatives, are presented in Figures 5.5 and 5.6 respectively. The accuracy of these results is found to be comparable to that observed above, demonstrating the flexibility provided by the rational-function model. Because it seems impractical to measure the frequency derivatives of EM observables, we note that a model based on using only frequency samples, akin to the results of Figure 5.6, is how experimental data would have to be employed. Even then, only two samples per resonance are needed, at a maximum, but fewer might be required if the effects of the resonances on the $j\omega$ axis of the complex-frequency plane are relatively smooth.

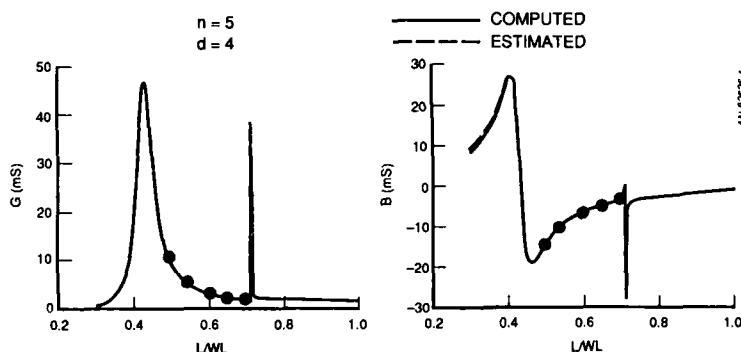


Figure 5.5. Input admittance of a forked monopole on a perfectly conducting ground. Directly computed values (_____) are compared with the rational-function model (_____) employing $n = 5$ and $d = 4$ based on the antenna admittance and its first derivative at the five frequencies indicated by dots.

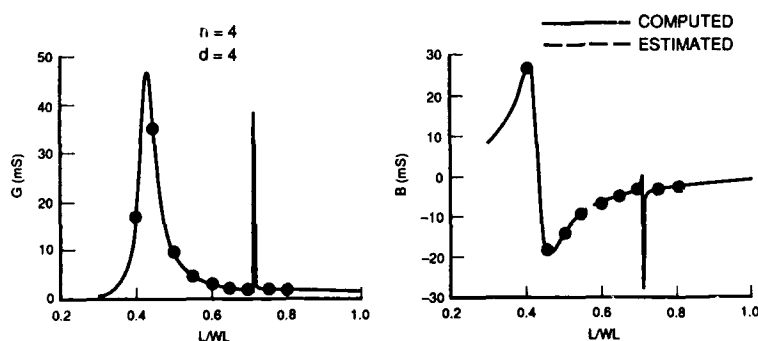


Figure 5.6. Input admittance of a forked monopole on a perfectly conducting ground. Directly computed values (_____) are compared with the rational-function model (_____) employing $n = 4$ and $d = 4$ based on the antenna admittance at the nine frequencies indicated by dots.

6 SOME POSSIBLE EXTENSIONS USING EXPONENTIALS AND POLES

Exponential- and pole-series expansions have electromagnetics applications beyond those thusfar discussed. Two in the area of spectral-domain estimation are briefly considered below.

6.1 An Admittance-Matrix Function

The impedance and admittance matrices that arise in the moment method may be interpreted as N-port-network approximations to the problem being modeled. When only a single port is excited, as is the case for most antenna applications, then the current vector that results gives the self and mutual admittances for that port. Exciting the other N-1 ports one at a time similarly provides the N(N-1) remaining admittances. Since typical problems rarely require all N² admittances, the inverse matrix is usually not computed and the solution is developed by LU decomposition, or iteration. Whatever excitation is used and however the solution is developed, the procedures outlined above can provide a rational-function approximation for the computed response.

A logical generalization of this somewhat limited approach is to represent some, or all, of the admittance-matrix coefficients in rational-function form. This is most easily done by either computing the inverse of the impedance matrix if all N² matrix coefficients are desired, or alternatively, computing current vectors caused by a sequence of single-port excitations. Whichever approach is used, we observe that the admittance matrix can be written in the general form

$$\vec{Y}(s) = \begin{bmatrix} X & \cdots & X & y_{1a}(s) & \cdots & y_{1b}(s) & \cdots \\ X & \cdots & X & y_{2a}(s) & \cdots & y_{2b}(s) & \cdots \\ X & \cdots & X & y_{3a}(s) & \cdots & y_{3b}(s) & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ X & \cdots & X & y_{Na}(s) & \cdots & y_{Nb}(s) & \cdots \end{bmatrix} \quad (6.1a)$$

where the X's denote the admittances that have not been computed, and the $y_a(s)$ are those whose rational-function forms are modeled. Upon observing that the system poles are analytically the same whatever system observable has been used to obtain them, Eq. (6.1a) can be rewritten as

$$\vec{Y}(s) = \frac{1}{D(s)} \begin{bmatrix} X & \cdots & X & n_{1a}(s) & \cdots & n_{1b}(s) & \cdots \\ X & \cdots & X & n_{2a}(s) & \cdots & n_{2b}(s) & \cdots \\ X & \cdots & X & n_{3a}(s) & \cdots & n_{3b}(s) & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ X & \cdots & X & n_{Na}(s) & \cdots & n_{Nb}(s) & \cdots \end{bmatrix} \quad (6.1b)$$

where $1/D(s)$ is the common denominator polynomial, and the $n_a(s)$ are the numerator polynomials for each computed admittance. These would usually be found as column vectors since for single-port excitation, solution for the spatial current yields a column of admittances.

The validity of Eq. (6.1b) derives from the fact that the mode amplitudes in Eqs. (2.1) and (2.2) are the only parts of the pole response which are excitation dependent while the poles themselves are determined solely by the properties of the object for which $F(s)$ has been obtained. Observe that given part or all of the admittance matrix in the form of Eq. (6.1b), computing transient responses can be made much more efficient and convenient.

6.2 Adaptive Frequency Sampling

We mentioned above that application of the MBPE approach is expected to work best when the polynomials are of relatively low order, which means that only a small number of resonances can be spanned by a given model. It therefore follows that two or more separate rational-function models would be needed to cover a frequency interval spanning many object poles. While that may appear to be a disadvantage, it does provide the opportunity of developing an error measure and establishing a strategy for selecting the frequency points to be sampled.

This can be illustrated conceptually by considering two models $M_1(s)$ and $M_2(s)$ developed from samples f_1, \dots, f_3 and f_3, \dots, f_5 , respectively, using a frequency-sampled model. In their region of overlap, i.e., f_3 to f_5 , we should expect that some measure of their mismatch, say $\Delta M(s)$, could be used to mutually validate both models. This should also provide an indication about whether more samples are needed in that region (when $\Delta M(s) > E$), or whether the sampling interval should be increased (when $\Delta M(s) < E$), where E is a maximum acceptable error. A similar approach could be developed using the frequency-derivative model. The goal in either case would be to automate selection of the frequency samples of the response or its derivatives are computed or measured while ensuring that the rational-function model which results satisfies some specified error criterion.

An indication of the possibility for automating the sampling process to achieve some desired overall accuracy while minimizing the number of samples required is demonstrated in Figure 6.1. Presented there are the magnitudes of the difference between various models as a function of frequency. These results show that the difference between two models which use either different data samples (at least one) and/or are of different order, can indicate where the model error exceeds an acceptable threshold and consequently where another sample (or more) might be needed. A reliable way to estimate the model error is of course an essential property for a process to be made adaptive, to thereby develop a sampling strategy whose goal is to maximize the information provided by each new sample.

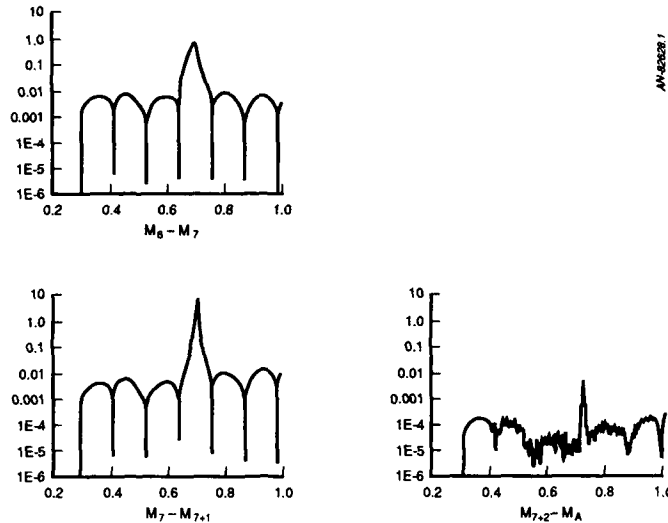


Figure 6.1. Adaptive frequency sampling demonstrated using MBPE. These results show the difference between models M_8 and M_7 using six ($n = 3, d = 2$) and seven ($n = 3, d = 3$) frequency samples respectively in part (a); the difference between model M_8 and an eight-sample model M_{7+1} ($n = 4, d = 3$) with the additional sample added near the error peak in part (b); and the difference between a nine-sample model M_{7+2} ($n = 4, d = 4$) with another sample added in the error peak and the actual results M_A in part (c). We see that an error estimate based on two rational-function models indicates where more sampling must be done in order to reduce the actual error to an acceptable value.

7 MBPE USING OTHER MODELS IN COMPUTATIONAL ELECTROMAGNETICS

Because of their ubiquity in electromagnetics, we have concentrated thusfar on MBPE applications which employ models based on exponential and pole series. Other applications using different models are also feasible. One is reducing the computer time required for filling a moment-method impedance matrix by replacing a rigorous Green's function with an MBPE approximation. Two examples are included here.

7.1 The Interface Problem

One approach to analyzing the fields of a source located near an infinite, planar interface which separates two half spaces is a Fourier expansion which leads to infinite integrals. These Sommerfeld integrals can be numerically time-consuming to evaluate, thus reducing the range of boundary-value problems involving interfaces which are practically modelable. An approach that has been found to greatly improve the efficiency (a factor of 100 or more) of obtaining the Sommerfeld fields needed for a moment-method model is MBPE. The model in this case might be based on simple bivariate interpolation for the case when the source (s') and field (s) points are on the same sides of the interface because the fields are then functions of only two variable, the lateral separation of s and s' , and the sum of their distances from the interface. In figure 7.1 is given a representative plot of the Sommerfeld fields for this kind of problem, the simplicity of which demonstrates the feasibility of replacing the rigorous field expressions with such a simple model.

When the more general problem where s and s' are separated by the interface is of interest, simple interpolation is no longer as practical and a more appropriate model must be employed. It has been demonstrated elsewhere [1] that models using various asymptotic and other closed-form approximations of the Sommerfeld integrals are then suitable. The models found to be most suitable might involve up to 24 (or more) terms each of which accounts for various propagation paths between s and s' , and whose amplitudes (the parameters) are obtained by matching the model to actual Sommerfeld field values. The result is an essentially rigorous model for the general interface problem for which the fill time of the moment-method impedance matrix is a factor of only four to eight times over that needed for the same object located in an infinite medium and which may be 100 times faster than directly employing the Sommerfeld integrals for the field computation.

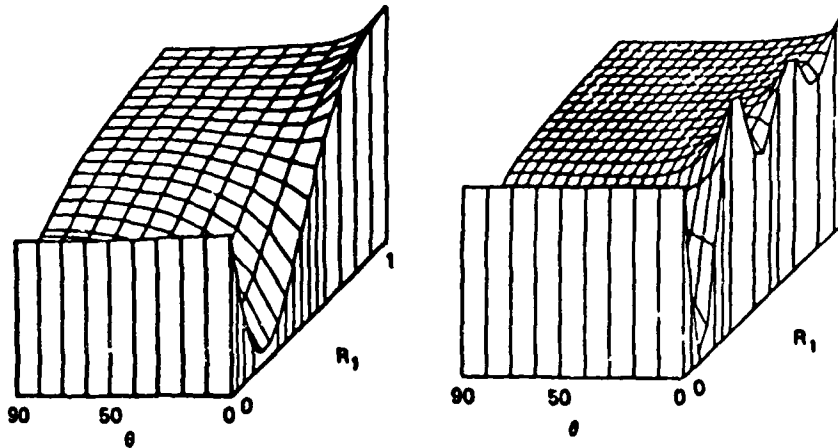


Figure 7.1. Mathematical complexity can sometimes obfuscate physical simplicity. While outwardly complex, the Sommerfeld integrals actually describe fairly simple physical phenomena, the fields of point current sources located near an infinite, planar interface, as illustrated in this plot for the one-sided problem (source and observer on the same side of the interface). It is consequently possible to replace the expensive Sommerfeld fields with simpler representations to achieve a savings in the impedance-fill time for a moment-method model of a factor of 100 or more.

7.2 Green's Function for Parallel-Plate Waveguide

Another Green's-function-type problem amenable to MBPE is that of point current sources located between infinite, parallel, perfectly conducting plates the field expressions for which involve an infinite series of images. For a rectangular waveguide or rectangular cavity, the corresponding expressions become doubly and triply infinite series, respectively. These series are generally very slowly converging, in part because the sequential terms can alternate in sign, making their direct computation from the defining expressions very time consuming. But, as for the case of the Sommerfeld problem just discussed, the fields given by these series, and which are needed for a moment-method model, are quite well-behaved as demonstrated in Figure 7.2. By using a simple model based on linear regression to replace the series-field expressions in the kernel of an appropriate integral equation, a speedup of 30 times has been achieved in modeling the a wire located in a parallel-plate waveguide.

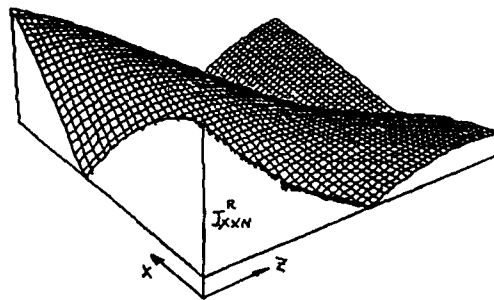


Figure 7.2. MBPE applied to the problem of the parallel-plate waveguide. The Green's function for this geometry involves an infinite series of images whose numerical convergence can be slow, thus making a moment-method model computationally expensive. The rigorous series Green's function can however describe a rather smoothly varying field as this figure as this plot of one component of the field demonstrates, thus making practical its approximation by a more efficiently computable representation. For this particular problem, regression analysis has been found useful, leading to a time savings of approximately 30 for a simple dipole located in the waveguide.

CONCLUSIONS

In this paper, we have discussed a variety of issues involved in use of model-based parameter estimation (MBPE) in electromagnetics. Models based on exponential-series (waveform domain) and pole-series (spectral domain) were summarized and a variety of electromagnetics observables were demonstrated to exhibit such mathematical behavior. It was also demonstrated that the parameters of exponential- and poles-series models can be obtained from either function sampling or function-derivative sampling, leading in every case to solution of a linear system of equations whose coefficients are comprised of the sampled data.

Two main applications of MBPE were demonstrated for:

- 1) Increasing the efficiency of the basic model computation by replacing rigorous, but computationally costly, Green's functions that appear in an integral-equation model with approximate but simpler models whose computation is much faster; and
- 2) Increasing the utilization efficiency of the model computations that are performed by reducing the number of times they must be evaluated via using physically motivated models to approximate the results of making additional model computations.

The basic idea in either case is using samples of an electromagnetic quantity to compute the coefficients (parameters) of some appropriate simplified expression (the model) by which that quantity can be replaced. For Green's-function MBPE, a variety of model types might be employed depending on the details of the actual problem. For applications in the latter area, for example in determining an electromagnetic response over some specified frequency band, a model based on rational-function approximations is found to be useful because the resonance nature of such responses is thereby well represented. In this case, the data used for the parameter estimation can be based on a combination of spectral samples and spectral derivatives to achieve an additional time savings. This is because the derivative information can be computed in a time proportional to N^2 as contrasted with the N^3 dependency of solving the original problem.

It should be clear that MBPE could develop into an important tool for not only increasing the efficiency of model applications, but in yielding more physically useful representations of those models.

REFERENCES

- [1] Burke, G. J. and E. K. Miller (1984), "Modeling Antennas Near to and Penetrating a Lossy Interface", *IEEE Transactions on Antennas and Propagation*, AP-32, pp. 1040-1049.
- [2] Miller, E. K. (1987), "Model-Based Parameter-Estimation Applications in Electromagnetics", presented at NATO Advanced Study Institute on Modeling and Measurement in Electromagnetic Analysis and Synthesis Problems, Il Ciocco, Italy, to be published by Sijthoff and Noordhoff.
- [3] Prony, R. (1795), "Essai Experimental et Analytique sur les Loix de la Dilatabilité de Fluides Elastiques et sur Celles de la Force Expansive de la Vapeur de L'alkool, a Differentes Temperatures", *J. l'Ecole Polytech.* (Paris), 1, pp. 24-76.
- [4] Hildebrand, F. (1956), *Introduction to Numerical Analysis*, McGraw-Hill, New York, NY.
- [5] Kay, S. M., and S. L. Marple (1981), "Spectrum Analysis—A Modern Perspective", *IEEE Proceedings*, 69, no. 11.
- [6] Baum C. E. (1976), "The Singularity Expansion Method", Chapter 3 in *Transient Electromagnetic Fields*, Ed. L. F. Felsen, Springer-Verlag, New York.
- [7] Burke, G. J. and E. K. Miller (1988), "Use of Frequency-Derivative Information to Reconstruct an Electromagnetic Transfer Function", *Proceedings of the Fourth Annual ACES Review*, Naval Postgraduate School, Monterey, CA, March 22-24.
- [8] Brittingham, J. N., E. K. Miller and J. L. Willows (1980), "Pole Extraction from Real-Frequency Information", *Proceedings of the IEEE*, 68, pp. 263-273.
- [9] Jones, D. S. (1979), *Methods in Electromagnetic Wave Propagation*, Oxford University Press, Engineering Science Series, Oxford, United Kingdom.
- [10] Miller, E. K., G. M. Pjerrou, B. J. Maxum, G. J. Burke, S. Gee, D. E. Neely, A. J. Poggio, and A. R. Neureuther (1970), Final Technical Report on the Log-Periodic Scattering Array Program (U), MBAssociates, San Ramon, CA 94583, MB-R-70/105.
- [11] Burke, G. J. (1987), "Enhancements and Limitations of the Code NEC for Modeling Electrically Small Antennas", Lawrence Livermore National Laboratory, UCID-20970, January.
- [12] Van Blaricum, M. L., and R. Mittra (1975), "A Technique for Extracting the Poles and Residues of a System Directly from its Transient Response", *IEEE Trans. Antennas and Prop.*, 23, pp. 777-781.
- [13] Poggio, A. J., M. L. Van Blaricum, E. K. Miller, and R. Mittra (1978), "Evaluation of a Processing Technique for Transient Data", *IEEE Trans. Antennas and Prop.*, 26, pp. 165-173.
- [14] Van Blaricum, M. L., and R. Mittra (1978), "Problems and Solutions Associated with Prony's Method for Processing Transient Data", *IEEE Trans. Antennas and Prop.*, 26, pp. 174-182.
- [15] E. K. Miller (1981), "Natural-Mode Methods in Frequency- and Time-Domain Analysis", in *Theoretical Methods for Determining the Interaction of Electromagnetic Waves with Structures*, Sijthoff and Noordhoff, pp. 173-212.

**Measurement and Computer Simulation of Antennas
on Ships and Aircraft for Results of Operational Reliability**

Stanley J. Kubina
EMC Laboratory, Concordia University
Montreal, Que. Canada

1. INTRODUCTION

The review of the status of Computational Electromagnetics by Miller[1] and the exposition by Burke[2] of the developments in one of the more important computer codes in the application of the electric field integral equation method, the Numerical Electromagnetic Code (NEC), coupled with Molinet's[3] summary of progress in techniques based on the Geometrical Theory of Diffraction (GTD), provide a clear perspective on the maturity of the modern discipline of Computational Electromagnetics and its potential. Audone's[4] exposition of the application to the computation of Radar Scattering Cross-section (RCS) is an indication of the breadth of practical applications and his exploitation of modern near-field measurement techniques reminds us of progress in the measurement discipline which is essential to the validation or calibration of computational modelling methodology when applied to complex structures such as aircraft and ships.

The title of this paper suggests a target standard for the confidence level of radiation patterns, be they measured or computed. If they are to be used for operational performance assessment purposes, they must be adequate for the determination of range or avionic system performance in a volumetric or three-dimensional sense to cater for operational scenarios where a multitude of aircraft pitch, roll or yaw attitudes will be encountered with respect to the corresponding emitter or receptor system. The bandwidth of many systems, say for frequency hopping purposes in battlefield communication scenarios or in electronic warfare situations, requires that the radiation pattern reliability apply over the entire operating band of airborne or shipborne systems. The complex geometry of modern aerospace systems and the use of materials of various conductivities, such as composite materials, stress the importance of full-scale static or flight test measurements for verification purposes. Scale-model, full-scale and flight test measurement techniques are reviewed in AGARDograph No.300[5] on the Determination of Antennae Patterns and Radar Reflection Characteristics of Aircraft.

The latter monograph also presents some comparison results with computational models. Some of the results[5] presented therein for scale model and flight measurements show some serious disagreements in the lobe structure which would require some detailed examination. This also applies to the radiation patterns[7] obtained by flight measurement compared with those obtained using wire-grid models and integral equation modelling methods. In cases like these, it is difficult to judge which is the more reliable result. It continues to be important that an extensive data base of results be sought of comparisons such as these where the circumstances of the measurements are known in detail and the model detail is sufficient for its replication by others.

While the comments made thus far are made from the perspective of final antenna installations on production aircraft, the life-cycle of aerospace systems gives rise to different levels of emphasis on antenna pattern information. In some cases comparisons are meaningful even though absolute accuracy is doubtful and measurements are either not possible or impractical because of time constraints on design decisions. Under these circumstances the reliability of computer models must be based on the judgment of the user and the precautions and guidelines which he has implemented. The fact that there are few examples in the literature of computer models for large complex structures being created and used 'ab initio', and subsequently corroborated fully by scale-model or full-scale measurements, suggests that more work remains to be done on code validation and model creation guidelines before the methodology can become a self-sufficient discipline. This is not to be interpreted as a recommendation against the use of computational techniques, but rather as a caution that the user exercise them with the same degree of care which is required for precise scale-model or full-scale measurements.

In the examples which follow, the first one relates to an attempt to match measurements results completely over the entire 2-30MHz HF range for antennas on a large patrol aircraft. The second example discusses the problem of validating computer models of HF antennas on a helicopter and using computer models to generate radiation pattern information which cannot be obtained by measurements. The third series of examples discusses the use of NEC computer models to analyze top-side ship

configurations where measurement results are not available and only self-validation measures are available or at best comparisons with an alternate GTD computer modelling technique. In one of these cases the special problem of source antenna modelling arises as an important issue. The last case involving the retrofit of a jammer antenna system in the reconfiguration program of an executive jet illustrates that computer modelling techniques can only be used to bound some aspects of the radiation problem and full-scale measurements must be used to obtain the operationally relevant radiation pattern and antenna isolation information.

2. HF ANTENNAS ON A LARGE PATROL AIRCRAFT

The Canadian CP-140/AURORA long range patrol aircraft program had led to the generation of an extensive radiation pattern data base both for low frequency and higher frequency antennas. Some of these data resulted due to its commonality with the P-3C/ORION configuration and the remainder was generated for the antenna locations which are unique to the CP-140, at the Rye Canyon Antenna Test Facility of Lockheed. It remains a unique resource for the testing of computer modelling methodology for both integral equation methods as well as GTD techniques.

Some of the results obtained by using NEC to model the airframe and the two wire antennas on the CP-140 had been discussed in AGARD Lecture Series 131[8] and later in the 1987 NATO Advanced Study Institute in Italy[9]. Only an up-dated summary is presented here to illustrate the process of optimizing a model to have an extensive bandwidth, in the case where scale-model measurement results are already available but which had been measured some time ago. The theoretical and algorithmic basis of NEC has been discussed elsewhere[4] and will not be repeated here. Special features such as the extended kernel, the junction constraints and source modelling are important features which effect the model creation guidelines.

The earlier work[8] showed that simple stick models of aircraft will provide approximate radiation pattern results over a narrow frequency range. When the models are to have considerable bandwidth and are required to produce reliable radiation patterns at the higher HF frequencies, complex models must be created. The creation of these models is best achieved by specialized computer-aided techniques which keep the user in touch with key parameters of the model which affect its performance. This comment applies to all cases of NEC modelling which are discussed.

2.1 THE MODEL CREATION ART

The user of the NEC code must observe certain guidelines in constructing a wire antenna model for analysis by the program. The guidelines are necessary so that the basic assumptions upon which NEC is based will not be violated.

The "thin wire" assumption requires that the wire diameter must be much less than the wavelength. In subdividing each wire into "segments", the length of each "segment" must be shorter than one-tenth wavelength, and segments shorter than one-twentieth are often used. Segments must not be shorter than one-thousandth of the wavelength. The segment should be much longer than the diameter of the wire. Segments longer than 8 times the wire radius are safe, and segments as short as the wire diameter are sometimes used. With the "extended kernel", segments should be longer than the wire diameter, although segments as short as half the wire radius are sometimes used. Parallel wires must be much farther apart than their diameter. Restrictions apply at wire junctions[10]. When wires of different radii are connected, the ratio of the larger to the smaller radius must be less than 5. The segment length must not be too dissimilar when two wires are connected. Differences larger than a factor of five are to be avoided. A more subtle restriction requires that the match point on a segment at a junction be located outside the volume of all the wires connected to the junction. Non-physical results can be seen if this restriction is violated.

These restrictions are local, relating segment length and wire radius within one wire, and on adjacent wires making up a junction. When the NEC program is used to analyse complex geometries in which a solid body of complex shape is replaced by a "grid" or mesh of wires describing its shape, the wires of the grid must satisfy the guidelines set out above.

The basic code requirements are that the start and end co-ordinates and the diameter of each wire be specified. For an aircraft model with 350 wires, some 600 vertices must be scaled and tabulated from the respective views of the aircraft drawings. Errors of overlap and discontinuity can frequently occur and these are not necessarily detected by the computer code itself. The computer-aided model creation system called DIDECE was designed to simplify and accelerate this process, to make it less error-prone and to add an interactive dynamic dimension to the exercise which makes the validity criteria and model features more visible to the engineer. The acronym DIDECE[11] means Digitize, Display, Edit and Convert.

The necessary first step in both the manual as well as the computer-aided process is the marking of the vertices on accurate dimensioned three-view drawings of the structure. The model designer then proceeds to "digitize" the vertices on a digitizing tablet. The drawing is first calibrated by digitizing two known reference points and then the co-ordinates of each vertex are determined sequentially: X,Y co-ordinates from a top-view and Y,Z co-ordinates from a side-view. Points can be linked automatically to form wires, or their linkage can be specified by listing their vertex numbers. At each stage of the digitization of any section of the structure, the vertices and wires can be displayed in multi-port displays with selected views corresponding to the source drawings and/or isometric 3-D displays.

Wires and vertices can be deleted or added as desired. For large symmetrical structures commands such as "REFLECT" and "MERGE" exploit symmetry and allow progressive build-up of a complex model from individual, manageably small sections. Colour-coded 3-D displays can show each wire in a colour that corresponds to a coded value of diameter or length. This provides meaningful visualization of these parameters as they relate to modelling criteria for variable sections of the model. If the model is found to be suitable, a final command will automatically format the internal files into an input data set that conforms to the format requirements of NEC or any other computer code that is being used. An additional off-line program called CHECK[8] will verify the modelling guidelines described above.

This modelling process was used to create the CP-140 model shown in Figure 2.1. It is shown alongside the antenna layout drawing for the CP-140. The D I D E C interactive model building system accelerates the process of developing wire-grid models by factors exceeding 20:1 over any manual process. It does so without mechanical measurement errors other than the tablet tolerances. These errors can be removed at will by numerical input. In the process of model creation it provides the medium for the creator to have better insight into the significant parameters that are important to the success of the model simulation. D I D E C is also available now for surface patch modelling purposes.

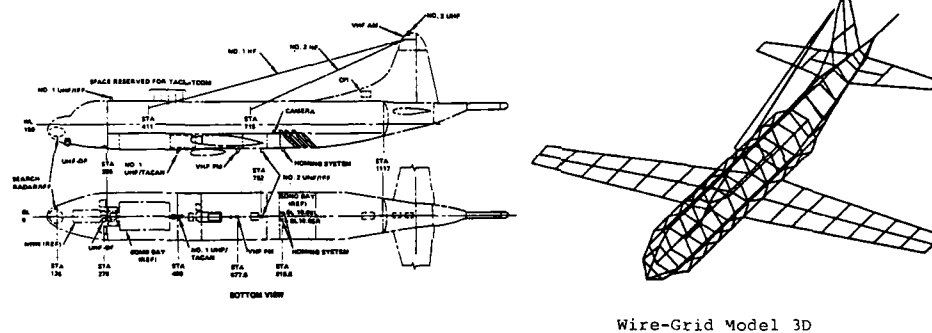


Figure 2.1 - P-3/CP-140 Antenna Arrangement & Wire-Grid Model

For reasons of economy, the model shown in the figure (3D) was intended to cover the 2-30MHz HF frequency band. It consists of 327 segments. The main portion of the fuselage has an octagonal section (i.e. 8 longitudinal wires) with the two upper longitudinal wires arranged to coincide with the lateral location of the feedpoints of the two wire antennas. The segmentation has been selected to have the segment lengths as 0.2λ at 30MHz. Thus the patterns would be expected to degrade at the higher frequencies. The M.A.D. boom on the tail is represented by a single wire.

In this first model, radii have been selected to produce the equivalent surface area in the computer model. In other related studies of model bandwidth [12] it was found that the best choice lies somewhere between a radius for equal area and that for equal perimeter. Segment radii are tapered on the wing and stabilizer surfaces as well as the non-uniform portion of the fuselage in order to minimize radius discontinuities.

What is desired from a computer model is that it produce radiation pattern data of sufficient accuracy to be useful for system performance calculations. As a second objective, it is also useful that the impedance of the antennas be calculated in order to anticipate tuning system requirements, although it is generally recognized that these calculations are less reliable than pattern calculations. With multi-antenna systems, it is also desired that the coupling between antennas be estimated.

2.2 RADIATION PATTERN RESULTS

In comparisons of radiation patterns over four octaves, the analyst is overwhelmed by the mass of pattern information that must be evaluated in order to properly sample the performance of the model on a volumetric basis over the frequency band. Unfortunately, at this time, there is no possibility to produce current distribution measurements which would be directly comparable to the currents obtained from the NEC matrix solution. In a previous work[9] single number parameters were discussed which facilitate these comparisons. Thus terms such as radiation pattern efficiency and percent E-theta are used. The first is the ratio of the amount of power in a 'useful' sector(± 30 degrees from the horizon), and the second is the ratio of power in the vertical component to the total radiated power. These parameters are directly related to communications system performance. Percent E-theta values were also available from the experimental data set, and these also are compared.

In previous work[8,9], the results obtained with model 3D were presented over the entire 2-30MHz range. There was surprisingly good agreement in lobe and null structure and null locations except at 24MHz. This is shown in Figure 2.2. With complex models it is difficult to do model optimization in a systematic way because of the number of variables that can be perturbed and the difficulty of relating these variations to canonical examples. However model refinement was carried out by reducing the radii by 5/12 to have closer agreement with segment guidelines at higher frequencies and by tapering the segments at junctions with large steps in radii. In addition the segmentation at the feed-points of the antennas was made more dense. Figure 2.3 also shows the results with the latest model(3J-364 segments) at 24MHz for the port antenna. It is seen that the lobe structure is better defined. The other radiation patterns remained essentially the same.

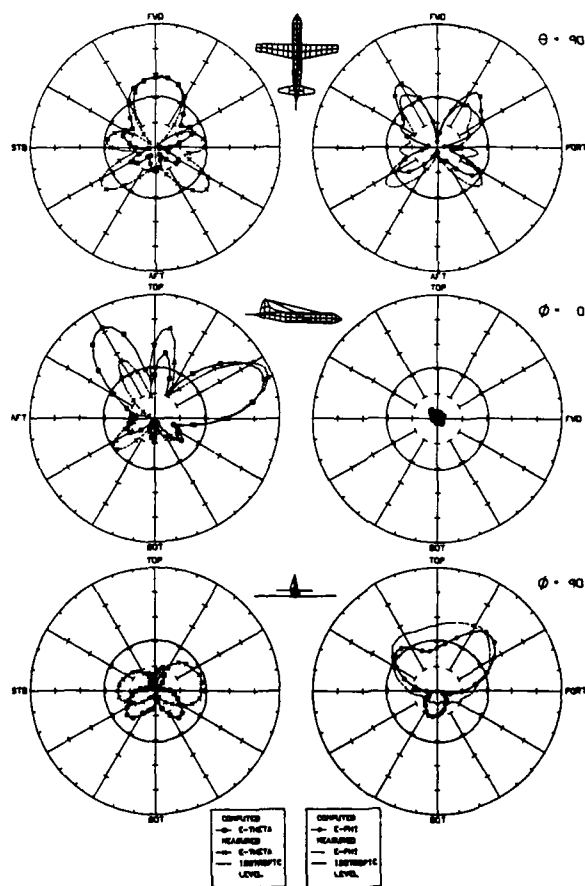


Figure 2.2 - Computed & Measured Radiation Patterns
Model 3D, 24 MHz

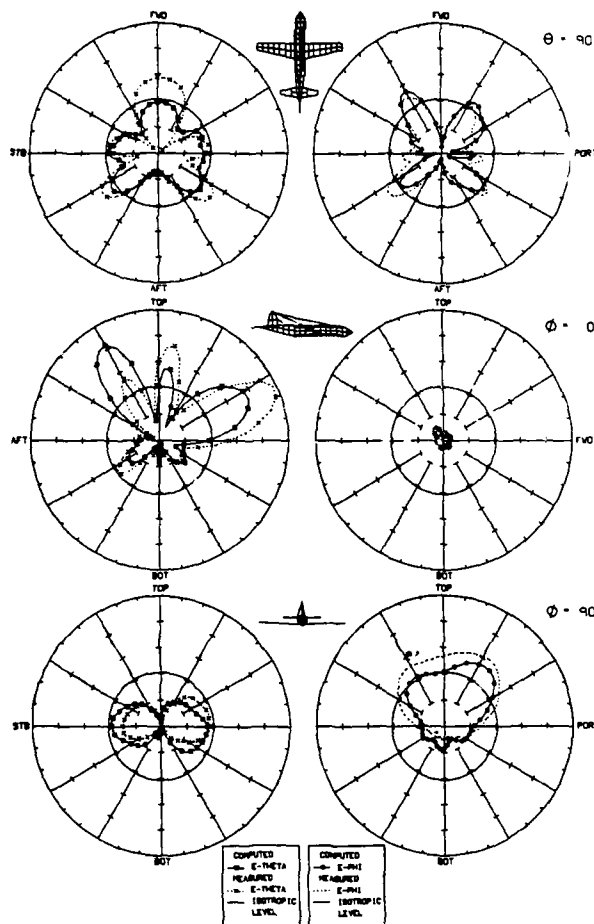


Figure 2.3 - Computed & Measured Radiation Patterns
Model 3J, 24 MHz

Judgments of radiation patterns based on principal plane patterns alone are difficult and often inconclusive. It has become standard practice to also use volumetric pattern plots such as those shown in Figure 2.4 for model 3J at 24MHz. Note that the peaks are fairly well reproduced. In addition to this format, the distorted sphere presentation shown in the next case study has been used. It allows a quicker evaluation of shape factors and an easier correlation with spatial angles, at the expense of more processing time.

The comparison of E -theta over the frequency band for model 3D is shown in Figure 2.5. The agreement is fairly good except at 6MHz where the radiation pattern agreement is inconsistent with this numerical value. The reason for the discrepancy is not known. The original investigators are no longer available to clarify this point.

2.3 IMPEDANCE VALUES

The impedance values which were computed by NEC for model 3D are shown plotted in Figure 2.6 in comparison with measured values on the full-scale aircraft. The exact measurement conditions are not known. The impedance variations are typical of an open-circuited transmission line, but the disagreement between measured and computed values is substantial. This discrepancy gave rise to many verifications that the effective lengths of antennas on the model and on the actual aircraft were the same. No discrepancy was found. At the same time, any reasonable model refinement did not produce any radical changes in computed impedance values.

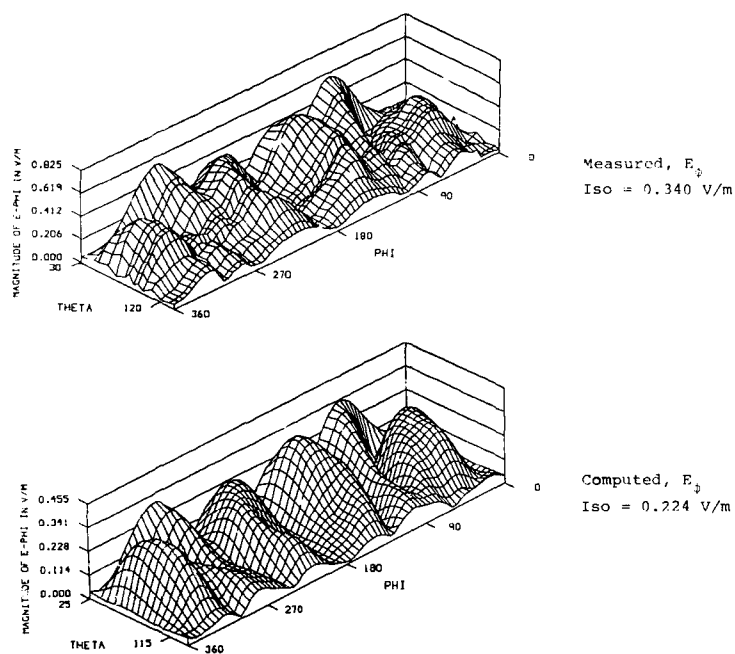


Figure 2.4 - Measured & Computed Volumetric Patterns,
Model 3J, 24 MHz

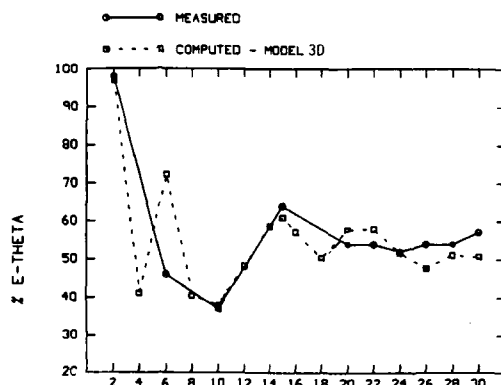


Figure 2.5 - Measured & Computed $\%F_0$ Curves,
Model 3D

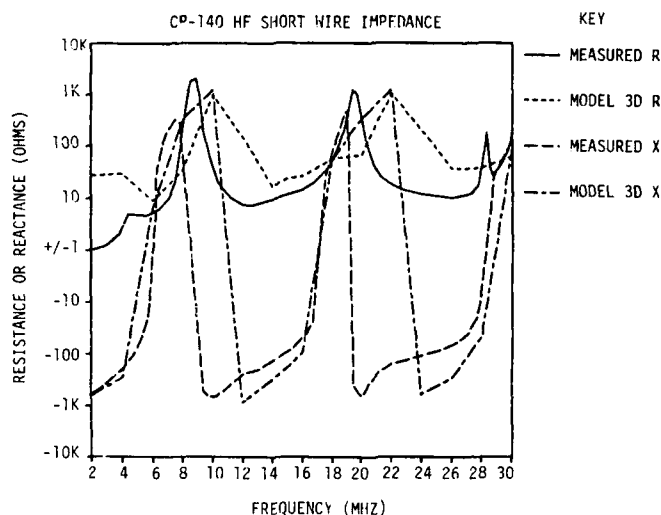


Figure 2.6 - Measured & Computed Impedance Curves, Model 3D, Starboard Antenna

Computations of impedance have received more attention in the recent literature[13] mainly through studies of model convergence and of the method of impedance calculation itself. In this study it was important to address the essential question as to what the NEC code actually calculates for the wire-grid model and what its relationship is to the real physical structure. It became apparent that the wire-grid model does not represent the shunting capacitance at the base of the feed-through insulator that is used on the aircraft. For the measurements on the actual aircraft it was also not clear whether the shunting lightning arrestor was in place. Based on these considerations, calculations at finer frequency steps were made over one frequency region and the calculated impedance values were shunted with discrete values of capacitance. Figure 2.7 illustrates the results of these corrections. Thus a shunting capacitance of 45pF produces values in reasonable agreement with the measured values. A portion of this shunting value can be accounted for by the measured base capacitance of the feed-through insulator and the lightning arrestor.

Other investigators[14] have applied a similar approach to calculations of terminal impedance by moment methods.

2.4 SUMMARY COMMENTS

The CP-140 example illustrates the importance of an experimental data set for model validation and optimization purposes. The Applied Computational Electromagnetics Society[15] is fostering the archiving of experimental and computational data sets which can be used by the community for the validation of computer codes and models. The dissemination of similar information within the AGARD technical community should serve a very useful and more application-oriented purpose.

The modelling guidelines which were applied to the NEC model optimization process did produce results which could reasonably be used in communication performance assessment and for the comparison of the performance of different antennas. Some of the minor discrepancies noted illustrate the usefulness of simultaneous programs of computational model development and scale-model measurements. The latter are not necessarily guaranteed to produce reliable results unless special care is taken with all the instrumentation details which guarantee their accuracy. This is illustrated in the following example.

In the modelling process used at the EMC Laboratory it has been found essential to have an integrated system of computer-aided graphics for model development and the analysis of the results. Insights into current distributions, for example, provide an important appreciation of model behaviour and its authenticity. Colour illustrations of this process are part of the oral presentation.

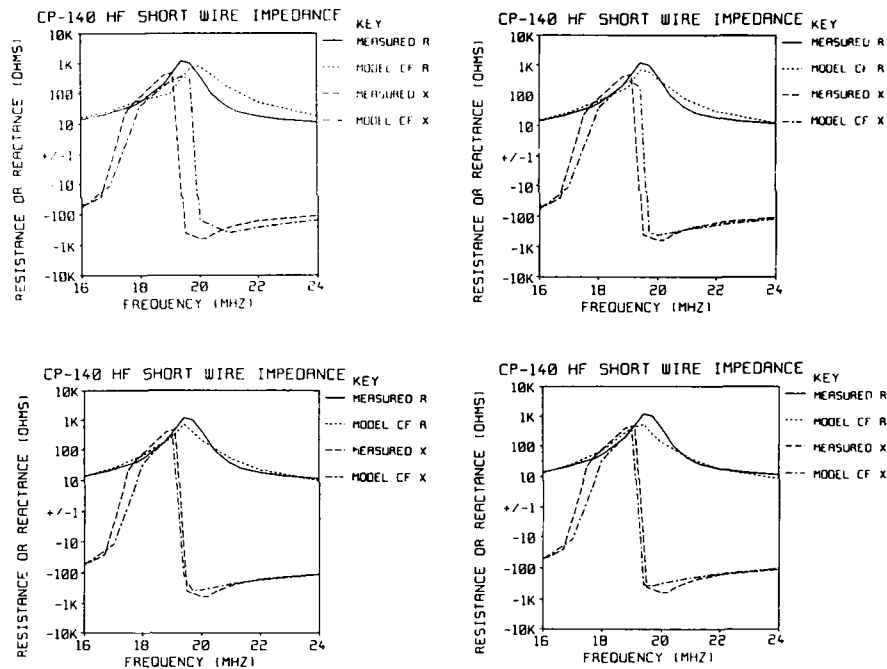


Figure 2.7 - Corrected Impedance Values, Model 3j

3. HELICOPTER HF ANTENNAS

Examples of the modelling of HF antennas on helicopters which had been presented before [8,9] focussed on special problems of resonance paths and rotor modulation. In the modelling of HF antennas on the CH-135/TWIN HUEY helicopter, the objective had been to continue the study of modelling guidelines further but also to arrive at models which could be used for the comparative evaluation of different antenna types. In addition, several actual examples of operational communication performance deficiencies were difficult to understand unless the propagation conditions were better understood in terms of field intensity variations with elevation angle and attitude as they would affect ground wave or nap-of-earth communication modes. Hence it was important to examine the radiation characteristics of helicopter antennas at typical flight altitudes.

It was fortuitous to have the computational model development proceed in parallel with a scale-model measurement program. The two techniques complemented each other surprisingly well and were useful in the identification of important experimental limitations. The two antennas which were analyzed were a shorted transmission line antenna (STL) and a Zig-Zag antenna, a name which derives from its geometry in order to extend its length.

3.1 RADIATION PATTERNS

Figure 3.1 shows the wire-grid models and the antennas plus the 1/24-th scale model on its model support tower. The models were developed using the same procedure with DIDEAC as described above. They have 302 and 305 segments respectively. The results of the wire-grid model analysis are presented over the entire 2-30MHz frequency range. Comparison with 1/24-scale model measurements showed some serious discrepancies between measured and computed results [16]. After extensive research [16] it was discovered that the model support rod used as shown in the figure was metallic and had a pronounced effect on the radiation patterns near 8MHz. It is doubtful that this would have been discovered had not the computational modelling taken place at the same time. The process demonstrates the utility and power of a combined computation/measurement approach for complex applications.

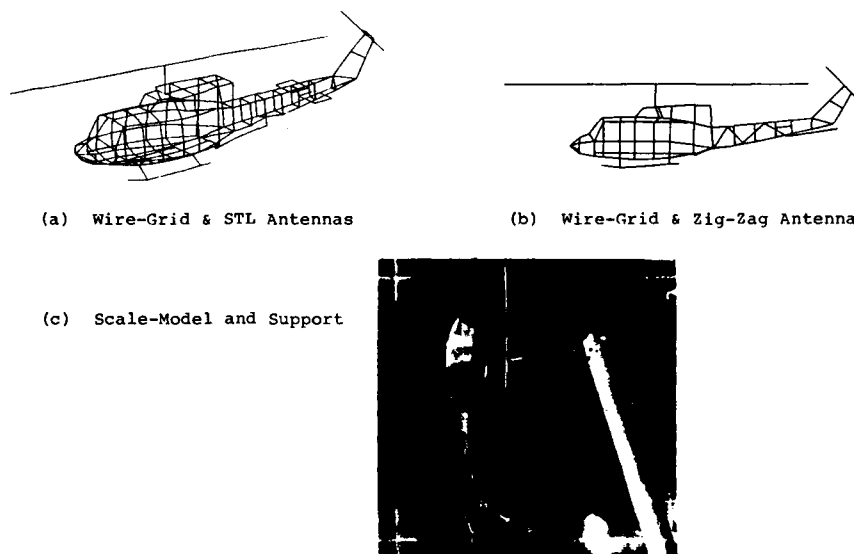


Figure 3.1 - Helicopter Models for Zig-Zag & Transmission Line Antennas

For a full operational analysis it is necessary to examine the volumetric radiation pattern characteristics. Again such computations were carried out at 2 MHz increments over the entire frequency band. Only the summary results are presented in this paper. Thus the radiation pattern efficiency for the two antennas versus frequency is shown in Figure 3.2(a). Note the difference between the two antennas at the higher frequencies. Such differences would have to be examined for their operational consequences.

For ground wave propagation, $\%E_{\theta}$ is important. Figure 3.2(b) plots this parameter versus frequency. Note the Zig-Zag antenna has higher values at 2 MHz. Should differences such as this be considered operationally significant, the coupler efficiency must be examined in order to establish that coupler efficiency differences will not cancel any pattern distribution advantages.

Having concentrated on the parameter values, it is always important to refer back to the actual pattern shapes for an appreciation of heading and attitude differences. For example, examining the plots of Figure 3.3 for the two antennas at 26 MHz and the distribution near the horizon, produces immediately the appreciation as to why the parameter values are so different. Thus for any special mission application, the parameters can be weighted based on the pattern detail that is of consequence.

The ease with which the complete pattern data can be generated from valid NEC models makes this procedure a valuable adjunct to HF system performance analysis.

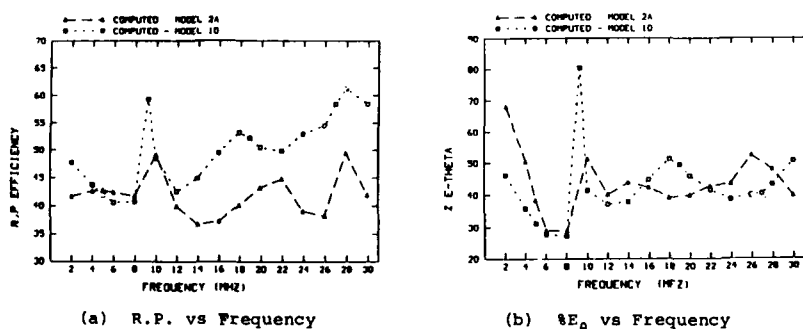


Figure 3.2 - Performance Parameters vs Frequency
Model 1D - STL Antenna
Model 2A - Zig-Zag Antenna

3.2 RADIATION PATTERNS VS ALTITUDE

Helicopters such as the CH-135 are used on many missions where the modes of operation of the HF communications system range from line-of-sight, ground wave to sky-wave propagation including high vertical incidence ionospheric propagation. These modes can be exercised while the aircraft is on the ground, hovering at low altitude, in transit altitudes or maneuvering at specific nominal altitudes. To identify the antenna contributions to these communications modes, it is necessary to know the radiation patterns under these conditions.

It has been normal practice to measure and compute radiation patterns under free-space conditions alone. Although limited full scale measurements have been carried out, they seldom include the vertical plane distribution. It would be possible to carry out such a measurement with scale models on a ship-range where booms allow vertical plane probing. However there is an obvious limitation to the range of altitudes that can be modelled.

The modelling of the ground plane in NEC allows the computation of the radiation patterns versus aircraft altitude. Such computations were carried for the STL antenna at 5 MHz with NEC for heights of 1 metre, 15, 30, 100 and 300 metres. The results are presented in two different formats, each of which brings out specific features of the resultant patterns.

Figures 3.4 and 3.5 show the vertical plane patterns. Here the formation of the lobe structure is clearly evident and instructive. It now allows the identification of the signal level at various elevation angles for ionospheric propagation purposes.

The three dimensional plots, shown in Figure 3.6 allow the complete conical set to be viewed and provide the missing information with regard to variations with elevation angle at various azimuth angles. Although the sampling is not dense enough for the lobes to be smoother, the nature of the pattern displays make the results easy to interpret by communications operators.

Figure 3.7 shows the same results for a representative value of ground conductivity. Note the difference in the lobe structure due to changed imaging properties of the ground plane. This degree of pattern information is difficult to obtain any other way.

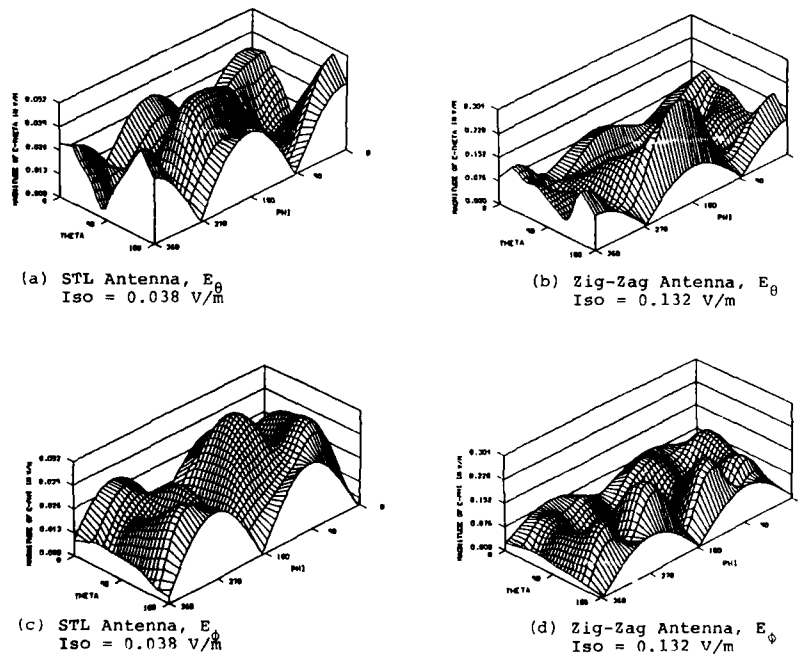


Figure 3.3 - Volumetric Patterns at 26 MHz

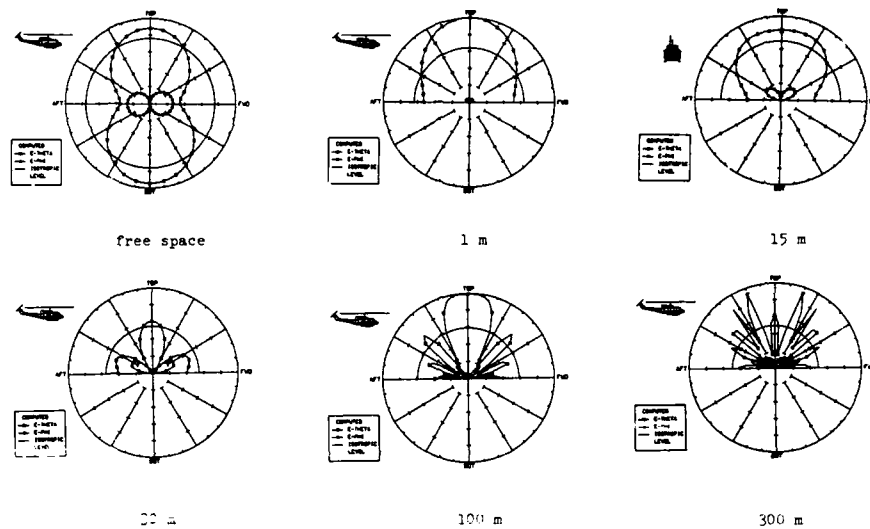


Figure 3.4 - Pattern Variation with Altitude
 $\phi = 0^\circ$, 5 MHz

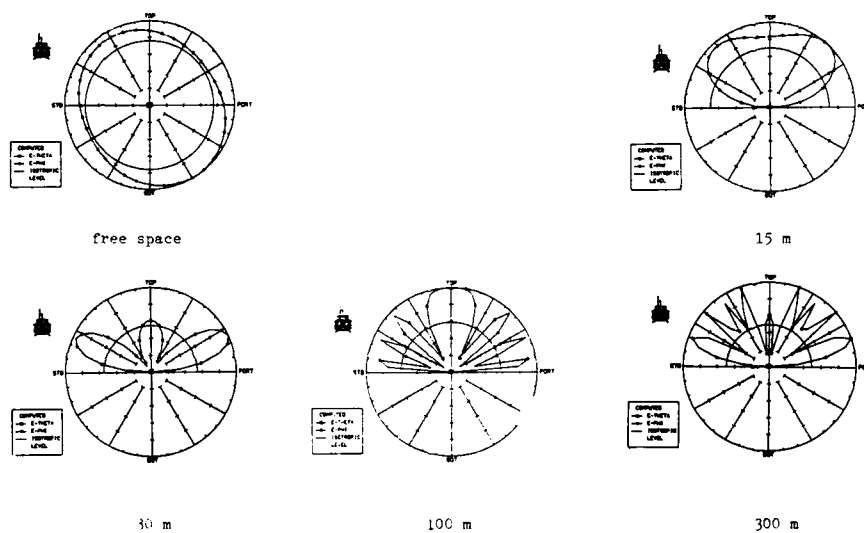


Figure 3.5 - Pattern Variation with Altitude
 $\phi = 90^\circ$, 5 MHz

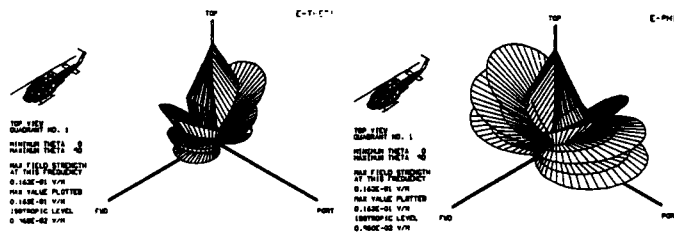
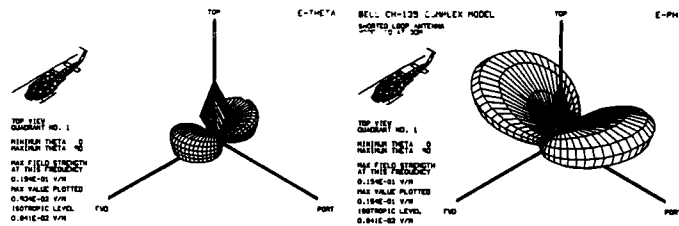
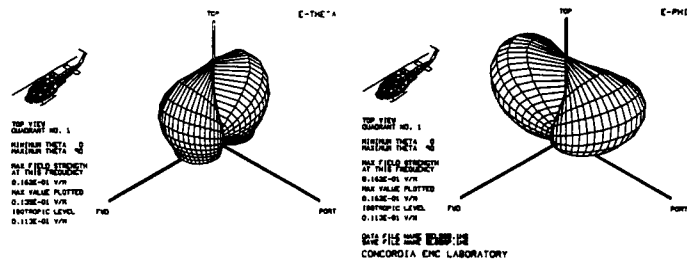
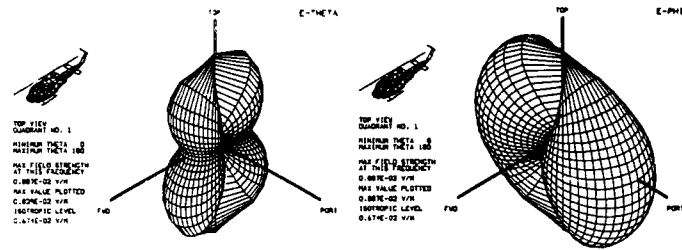
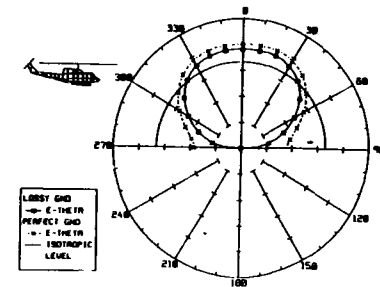
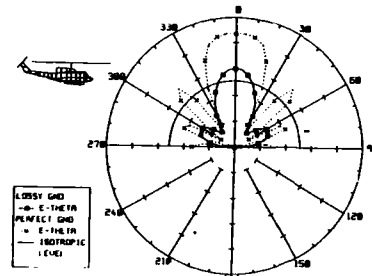


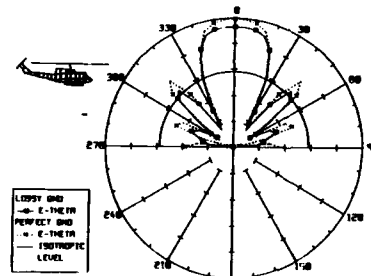
Figure 3.6 - STL Three-Dimensional Patterns (TDPAT)
vs Altitude
5 MHz, Perfect Ground



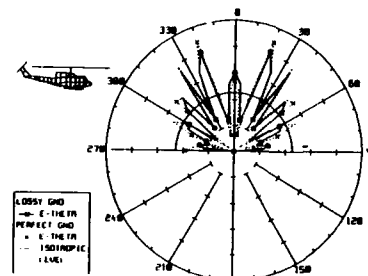
15 M



30 M



100 M



300 M

Figure 3.7 - Vertical Plane Patterns, 5 MHz
Ground Constants, $\sigma = 5 \text{ mS/m}$, $\epsilon_R = 15$,
E-theta, $\phi = 0^\circ$

3.3 CONCLUDING REMARKS

This helicopter case has shown how the NEC computer models can be used to compare the performance of two different antenna types on the CH-135 helicopter. The discrepancies between measurement results have been reconciled in terms of the specific measurement conditions because there was time to re-examine the measurement details. The use of the models to compute the change in radiation patterns with altitude permits a realistic analysis of HF system performance under a variety of flight profiles. Corresponding patterns under conditions of different ground conductivities are essential to the analysis of communications for realistic operational locations.

4. SHIPS' TOPSIDE DESIGN

Three cases of varying degrees of complexity have been selected to illustrate the computational modelling process where it is not possible to have access to measurements but the problem must be bounded. The first example of a UHF dipole and nearby mast[21] is quite tractable. The second example of a similar antenna near a scanner and mast becomes more complex to analyze while the third example of a TACAN antenna near a huge radar scanner introduces the problem of source modelling and analyzing the results in terms of its effect on system performance.

4.1 UHF ANTENNA AND MAST

Some of the compromises in ships' topside antenna design are discussed in the book by Preston E. Law[17]. For UHF communications antennas he mentions the mounting of dipoles on the ends of yardarms. In this first example, what is addressed is the dilemma faced by a project engineer who must approve the location of UHF antennas mounted on a yardarm near a sizeable tubular mast structure that supports other antenna systems and devices. Without quantitative data, say from brass scale-model measurements, he needs computations of pattern data in order to establish the extent of the perturbations which will occur. Some of the detail of an example which occurred during a ship modernization project is shown in Figure 4.1.

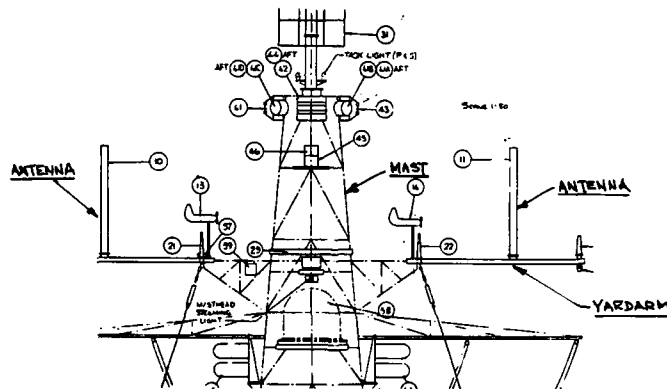


Figure 4.1 - Topside Detail - Rear View

The first impression in viewing the figure is the clutter of detail of other systems in addition to the two UHF antennas on the end of the yardarm and the main mast structure. The first question then is, what should be modelled and by what technique? In the absence of measurement data, it was prudent if not essential to have results from two different techniques for comparison purposes. Thus NEC and the Basic Scattering Code[18] were used.

The use of NEC at UHF frequencies for a structure of this size implies a large number of wire segments and costly execution times. This consideration led to the modelling of the mast alone rather than the complete modelling of the nearby environment including the less dominant structures. This approach was thought to be sufficient to determine the order of magnitude of the perturbation that would be produced in the inherent omnidirectional azimuth pattern of the UHF dipole antennas.

The corresponding GTD model was a box-like structure formed from a set of intersecting plates placed at the periphery of the mast shown in Figure 4.2. The plate model approximates the blockage and reflection effects of the structure while the joining edges provide the diffraction effects from the corners of the structure.

The mast model which was developed is shown in the four-ported DIDEC display of Figure 4.2. Although DIDEC allows the progressive development of individual portions of a structure, their concatenation, reflection and translation, this structure was too simple for the use of these powerful features. Rather as the details of the structure were identified in the drawings, the vertices were naturally scaled and tabulated by hand. However, the file manipulation features of DIDEC were too useful to ignore. Hence as the vertices were entered using the manual keyboard entry mode, the automatic wire segmentation feature was used to derive the three different models at the three specified frequencies of 230, 297 and 395 MHz.

Figure 4.3 shows the structure with the actual radii and segmentations that were used at 297 MHz. The segment lengths were chosen to be less than 0.2 and the radii were the estimated radii of the tubular members. The 685 segments of the 395 MHz frequency model represent a sizeable matrix size for NEC execution on the university CYBER 835.

For purposes of this study free-space calculations were sufficient to represent the perturbing effects of the mast, hence ground plane calculations were considered but deferred. Precise limits for allowable pattern distortion are difficult to estimate in complex cases such as this. The project engineer would be concerned, however, by 3dB pattern distortions occurring over a sizeable sector in azimuth.

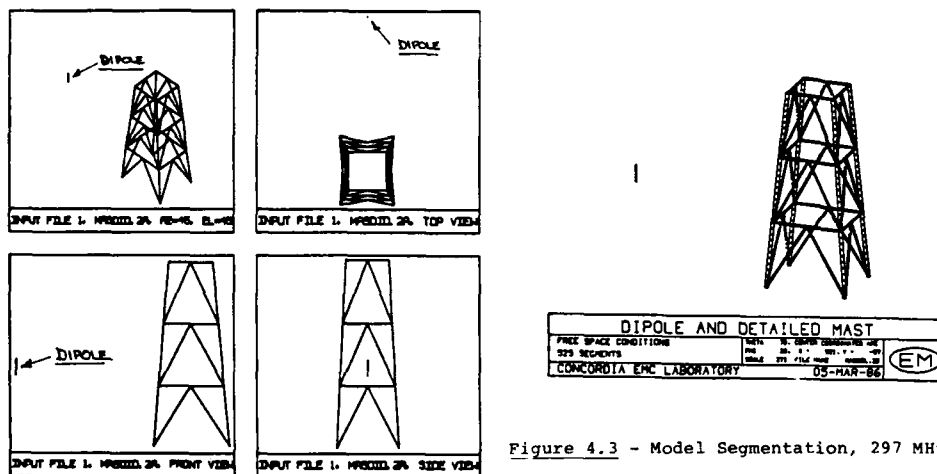


Figure 4.3 - Model Segmentation, 297 MHz

Figure 4.2 - DIDEC Display
Dipole and Mast

4.1.1 NEC RADIATION PATTERN CALCULATIONS

The results for the three frequencies are presented in Figures 4.4 through 4.6. The dominant features are the scalloping of the dipole patterns and the generation of a sizeable cross-polarized component. The null directly to port ranges from 8-14dB and the 3dB perturbation limits are exceeded over angular sectors ranging from 10 to 20 degrees. The three-dimensional displays show the extent of the disturbances in the elevation plane. Note that in these displays the scale is linear.

Azimuth Patterns

3D Pattern Displays

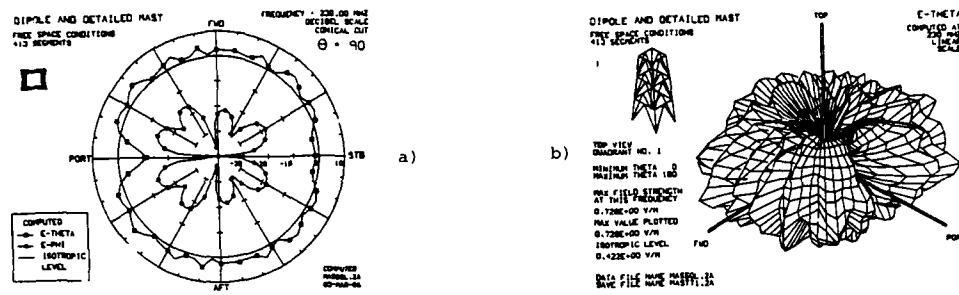


Figure 4.4 - Radiation Patterns, 230 MHz

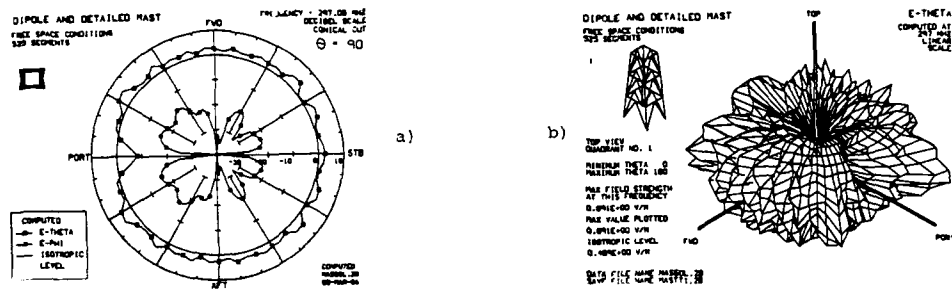


Figure 4.5 - Radiation Patterns, 297 MHz

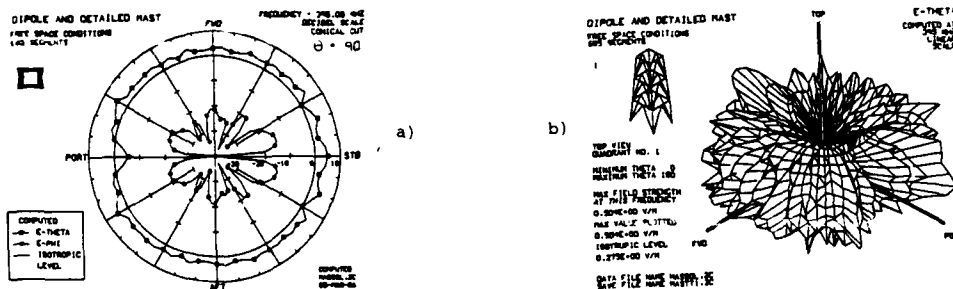
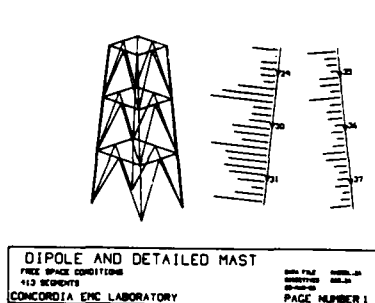


Figure 4.6 - Radiation Patterns, 395 MHz

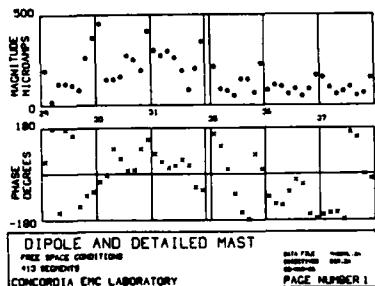
Are these results credible? To formulate an answer it is the practice at the EMC laboratory to examine the current distributions for self-consistency in a series of colour-coded displays with a computer code called SPECTRUM[19]. Although these cannot be reproduced in this paper, the reader can visualize that they would produce an appreciation of amplitude and phase distribution which could be linked to the radiation pattern results. Relatively large values of current are induced on the diagonal elements which, together with those on the horizontal members, give rise to the cross-polarized radiation component.

The correlation of these colour displays with the more conventional displays of amplitude and phase helps to train the analyst's eye in the appropriate interpretation of these results. An illustration of this is shown in Figure 4.7. The first part of this figure shows the amplitude differences between the front and rear legs of the tower at 230 MHz. Often it is meaningful to plot the amplitude and phase of the current versus segment number as shown in the lower portion of the figure. This is particularly useful in cases where it is important to identify resonant paths.

The examination of the radiation patterns and the subsequent analysis of the current distributions for an appreciation of their specific correlation in terms of current location, amplitude and phase, produces a basic level of confidence in the modelling and serves as an appropriate base for the comparison with GTD results.



a) Amplitude - Front, Rear Legs



b) Amplitude/Phase vs. Segment Number

Figure 4.7 - Alternate Current Displays

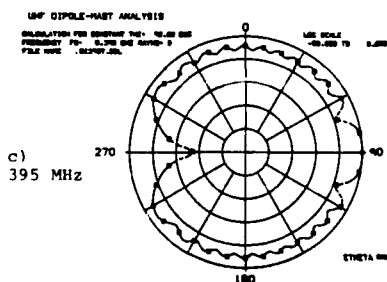
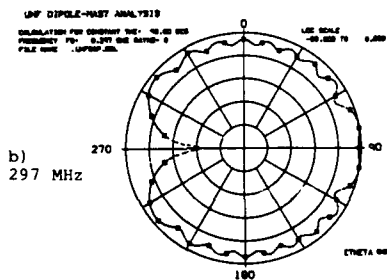
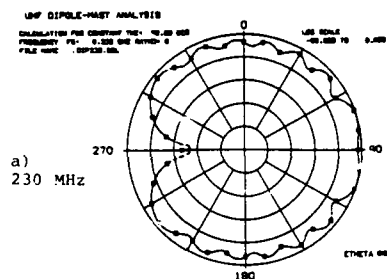


Figure 4.8 - GTD Results

4.1.1 RESULTS OF GTD MODELLING

The results from the Basic Scattering Code using the simplified box-like model is shown in Figure 4.8. It can be seen that except for the sharper nulls, the patterns are in good general agreement with the NEC calculations. As expected, there is considerable similarity between these patterns and the patterns of canonical dipole/plate reflector geometries[20]. This latter geometry has been the subject of many measurements in the past which show agreement with the referenced results. Note the cross-polarized component is zero in the GTD model.

The GTD computations are relatively inexpensive by comparison to those for the wire-grid model. In this regard, they are more suitable for the exploration of the distortion effects at finer frequency increments over the UHF range.

This example has established the pattern of modelling of the more complex cases which follow.

4.2 UHF ANTENNA MAST AND RADAR SCANNER

Another geometry that was analyzed was that shown in the DIDEDEC display of Figure 4.9. The arrow indicates the position of the UHF dipole with regard to the dominant structure of mast and scanner. Because of the larger size of this problem (>900 segments), it was only analyzed at 297 MHz. The radar scanner was modelled in the two extreme positions. Figure 4.10 shows the principal plane patterns for the case of the radar pointing to the side and Figure 4.11 shows them for the radar scanner pointing ahead. Once more the blockage effects and the generation of a cross-polarized component are evident. This class of problem becomes somewhat more unwieldy for GTD approximation, but sufficient modelling was done to corroborate the main features of the pattern. It should be pointed out that the outcome was to seek an alternate location for the antennas.

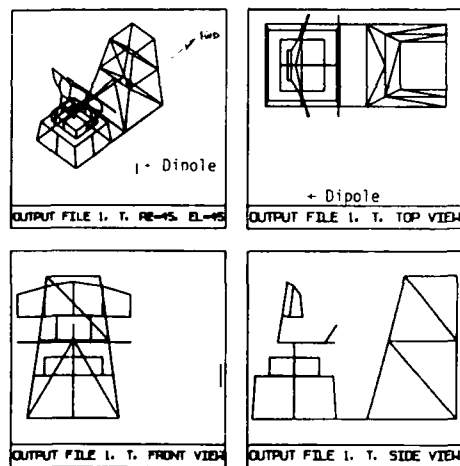


Figure 4.9 - DIDEDEC Display - UHF Antenna, Scanner and Mast

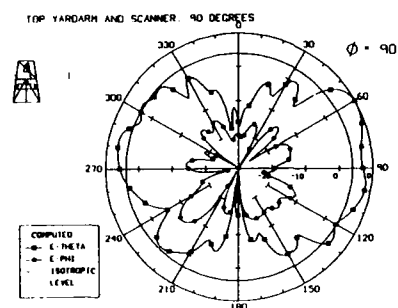
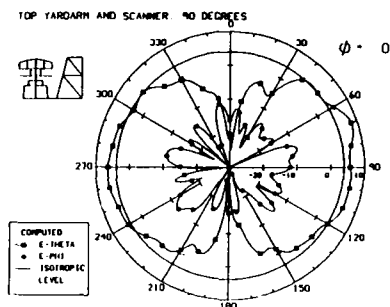
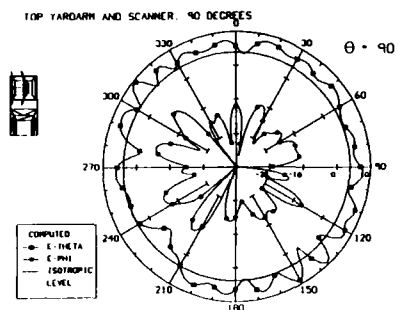


Figure 4.10 - Principal Plane
Patterns:
Dipole, Scanner,
and Mast
297 MHz
Scanner to Side

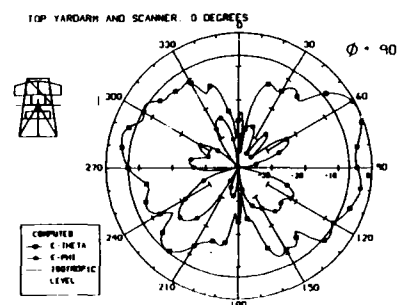
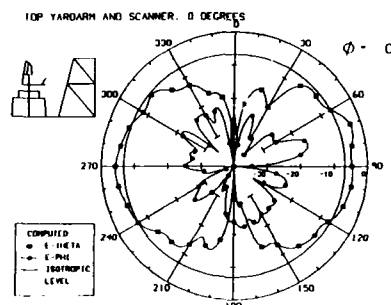
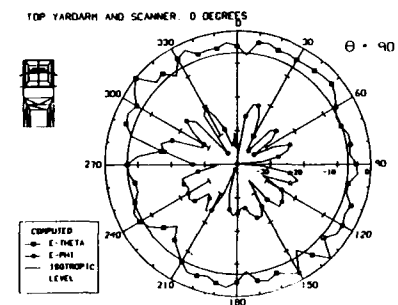


Figure 4.11 - Principal Plane
Patterns:
Dipole, Scanner,
and Mast
297 MHz
Scanner Forward

4.3 TACAN ANTENNA AND SCANNER

The most challenging case arose from an attempt to analyze the performance of a ground station TACAN antenna on the ship, in the presence of a large radar scanner as shown in Figure 4.12. The moving up to the 1GHz range makes this problem a better candidate for GTD techniques, except that the source modelling problem plus the structural complexity could not be reconciled with any available computer codes. The decision was to attempt a NEC modelling in progressive stages and try to quantify the scanner effects to a sufficient degree that the project engineer would have some basis for discussion at a pending design review.

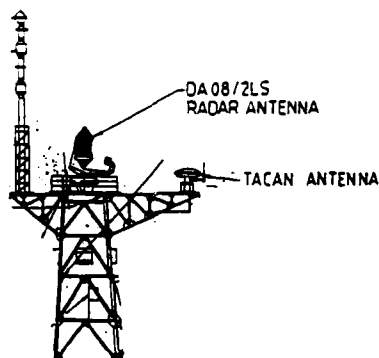


Figure 4.12 - TACAN Antenna and Scanner Installation

4.3.1 MODELLING THE TACAN ANTENNA AND SCANNER

Fundamental to the modelling process was the description of the TACAN antenna system which was sought in standard references[21]. The TACAN antenna array generates a scalloped cardioid pattern as shown in Figure 4.13. This pattern rotates at 15 revolutions per second as shown in Figure 4.13, the 15 cycle modulation providing the course and the 135 cycle providing the fine bearing information. As a first step this array had to be modelled to provide the required basic shape plus the percent modulation for the nine lobes. Systematic experimentation with NEC, based on the actual antenna geometry, led to the derivation of the array of active element and parasitics which produced the radiation pattern shown in the azimuth plane and in three-dimensional format in Figure 4.14. This array was used to illuminate the scanner in a number of representative positions.

A DIDECE-derived model of the scanner is shown in Figure 4.15. The distance from the center of the array to the scanner axis is approximately 18 feet. The scanner grid was relatively crude in order to keep the number of segments below 1000.

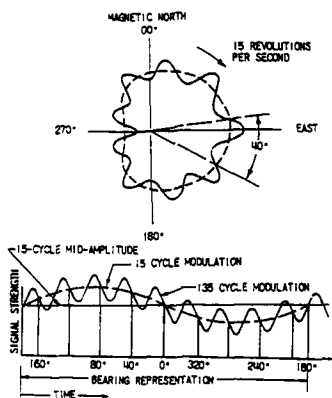


Figure 4.13 - TACAN Pattern Rotation for Course and Fine Bearing Determination

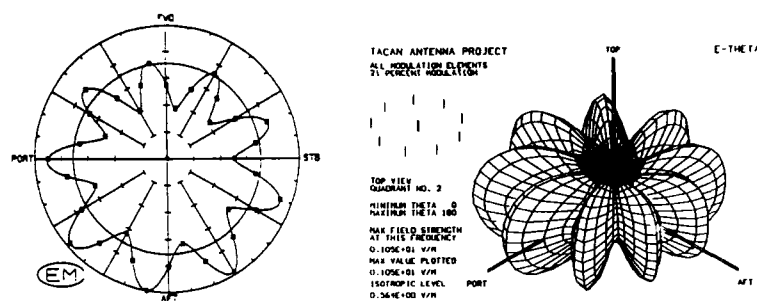


Figure 4.14 - Pattern of TACAN Antenna Model

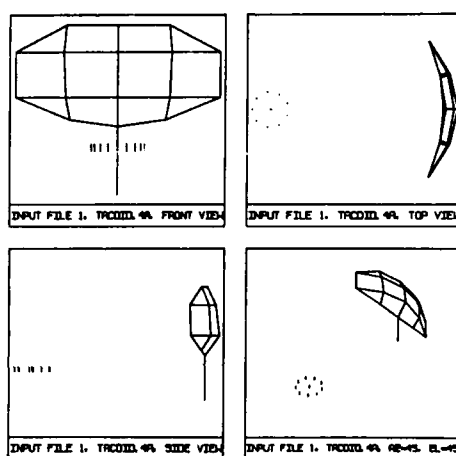


Figure 4.15 - DIDEC-Derived Model of Scanner

4.3.2 RADIATION PATTERNS VERSUS SCANNER POSITION

The radiation patterns were computed for the fundamental scanner positions as shown in Figure 4.16 and for two different TACAN antenna orientations with the scanner pointing forward-starboard as shown in Figure 4.17. It can be seen that there is considerable rippling of the basic TACAN antenna pattern. This information was passed on to the project director. However it was not clear how the TACAN bearing error would be affected. Clearly maximum distortion was observed for the forward-starboard scanner position.

Finally it was decided to sweep the TACAN antenna at one degree intervals over a 40 degree sector for the forward-starboard position of the scanner. From this information at any observer position in space the TACAN signal waveform was subjected to a least-squares analysis to determine the 15-cycle and 135-cycle waveform components. From this information the bearing error was estimated noting that to produce 1 degree of physical error, nine degree of phase shift error is required in the 135Hz waveform. Such an error curve is shown in Figure 4.18. It must be emphasized that this represents a massive amount of NEC computation and data processing.

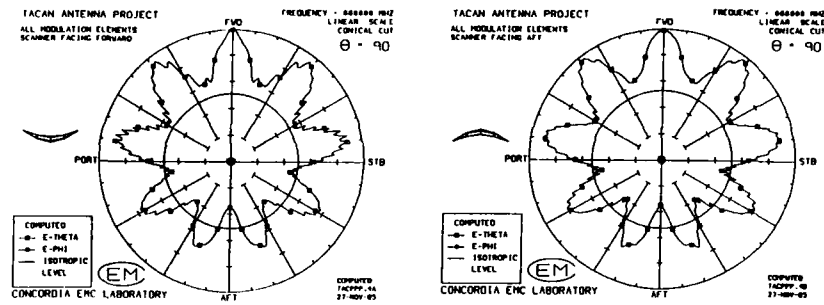


Figure 4.16 - TACAN Pattern vs Scanner Positions

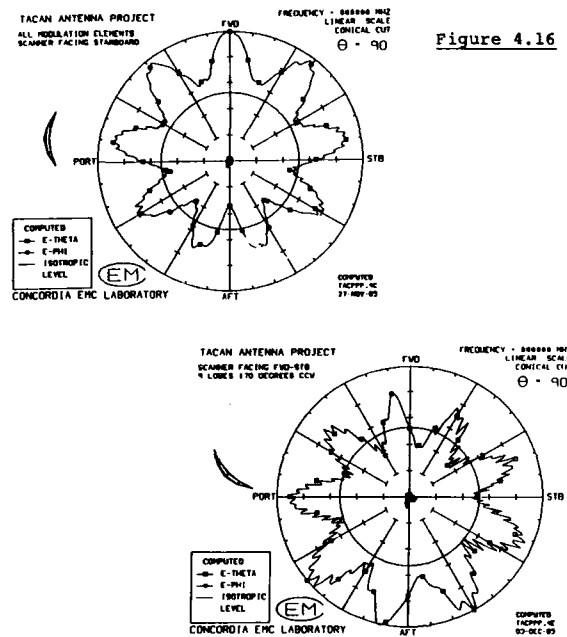


Figure 4.17 - TACAN 170° CCW and 10° CW for FWD/STBD Scanner Position

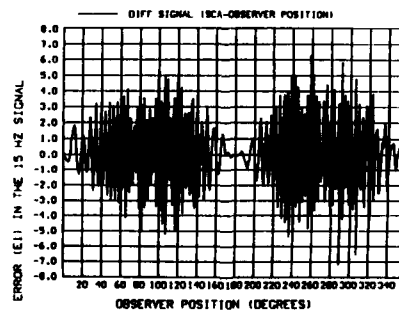


Figure 4.18 - TACAN Bearing Error vs Observer Position

4.4 CONCLUDING REMARKS

The ship's topside series of examples illustrates the complexity of the design problems and the need to have computational techniques available to guide the design optimization. It is hoped that large ship design projects can be used to generate supporting brass model data sets which can be used for validation of the approaches described above. The TACAN example illustrates the need in all computer codes to have flexibility in source representation in order to realistically emulate these practical systems.

5. ANTENNA LOCATIONS FOR AN I-BAND JAMMER

The Canadair CL-601 is an executive jet. The Canadian Forces undertook to implement a conversion to an Electronic Warfare (EW) training role. The antenna locations on the existing aircraft are shown in Figure 5.1. This last example restricts itself to a discussion of the main features of an antenna location study for an I-Band Jammer. Intelligent jammers use a transmit and receive antenna which must have a minimum of 60dB decoupling which had been established from a previous successful pod installation. The antennas are circularly polarized horns with beamwidths of approximately 25 and 40 degrees for the available antennas.

The task, simply stated, was to find a location for a rearward-facing antenna set and a forward-facing antenna set of transmitter and receiver antennas which would have the minimum level of coupling and result in installed antenna patterns with clear look angles, or a minimum of distortion of the inherent patterns of the horns. The options for a retrofit installation which would not involve extensive relocation of existing equipment are severely limited. Three major options seemed to be available. First the rearward facing antenna set could be mounted at the end of the tailcone. The forward-facing antenna set could consist of a transmitting antenna on top of the fuselage above the cockpit with a receiving antenna inside the radome, or both antennas could be mounted inside the radome. These are physical considerations only. Mounting of the antennas on the side of the nose was not practical in an inexpensive retrofit.

There are no available computer codes now which can model the source antennas completely on the oval conical nose sections or in the complexity of the nose installations with its bulkhead, radome and the weather radar scanner installation. However it was decided to try to bound the problem with the best approximations which could be made with the tools at hand and define the remainder of the information with a full-scale mockup measurement program.

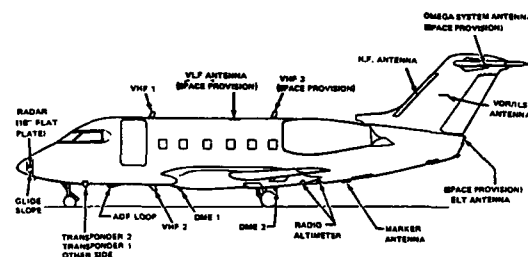


Figure 5.1 - Antenna Locations CL-601 Challenger

5.1 PRELIMINARY ANALYSIS

The EMC analysis code called Antenna-to-Antenna Propagation with Graphics or AAPG[23] has the capability to model aircraft as cylinders, cones and wing planes. In its computation of EMI margin values it calculates free space, creeping wave and edge diffraction losses for the dominant coupling paths between antennas. Model creation and execution are very rapid with interactive graphics. It was used to model the aft-facing cone as shown in Figure 5.2. This corroborated the insight that a side mounting of the horns would provide the necessary antenna isolation. Many other geometrical arrangements were studied in this way.

The top-mounted transmit antenna and the nose-mounted receive antenna combination was analyzed as shown in Figure 5.3. Again the results indicated that the antenna isolation would be achieved but that the radiation pattern coverage for each antenna location would have to be determined by some other means.

DESCRIPTION	HARMONIC NO 1 0.000 GHz
TRANSMITTER POWER	63.0
INTX CABLE LOSSES	0.0
INTX ANTENNA GAIN	-4.0
TRANSMISSION LOSS	-51.0
SURFACE SHADING	-10.1
EDGE SHADING	0.0
RCAR ANTENNA GAIN	-18.0
RCAR CABLE LOSSES	0.0
RCAR SENSITIVITY	-30.0
EMI MARGIN FIELD SHADING	0.0 \ 00.

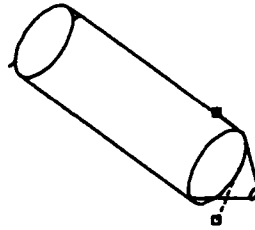


Figure 5.2 - AAPG Analysis of Rear-Facing Location

DESCRIPTION	HARMONIC NO 1 0.000 GHz
TRANSMITTER POWER	63.0
INTX CABLE LOSSES	0.0
INTX ANTENNA GAIN	0.0
TRANSMISSION LOSS	-60.1
SURFACE SHADING	-0.0
EDGE SHADING	0.0
RCAR ANTENNA GAIN	-18.0
RCAR CABLE LOSSES	0.0
RCAR SENSITIVITY	-30.0
EMI MARGIN FIELD SHADING	10.4 \ 00.

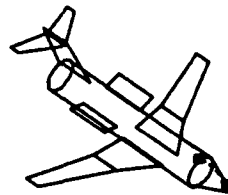


Figure 5.3 - AAPG Analysis of Forward Installation

The possibility of engineering an antenna installation entirely within the nose radome was attractive enough to pursue in detail. A basic ray clearance analysis was improvised using DIDECA as shown in Figure 5.4. Aircraft manufacturers now have more sophisticated CAD programs to establish clearances and view angles. However it was impossible to model the details of this installation with AAPG. A Basic Scattering Code Analysis was done with crossed-slots in the antenna locations in order to determine the reflection and blockage properties of the weather radar scanner plate array. This information was insufficient to make any serious judgments on the nature of the radiation patterns of the antennas mounted on the bulkhead. An accelerated full-scale mockup measurement program was arranged at the National Research Council facility in Ottawa in order to generate the radiation pattern and coupling information in order to meet the project deadlines.

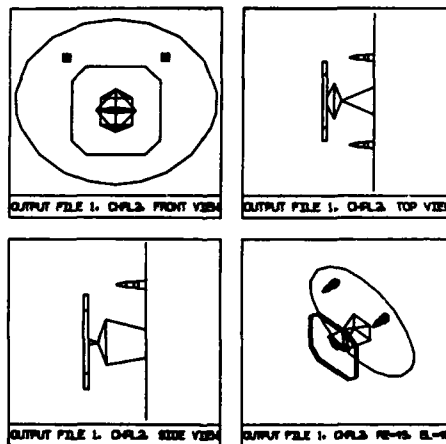


Figure 5.4 - DIDECA Diagram for Ray-Clearance Analysis

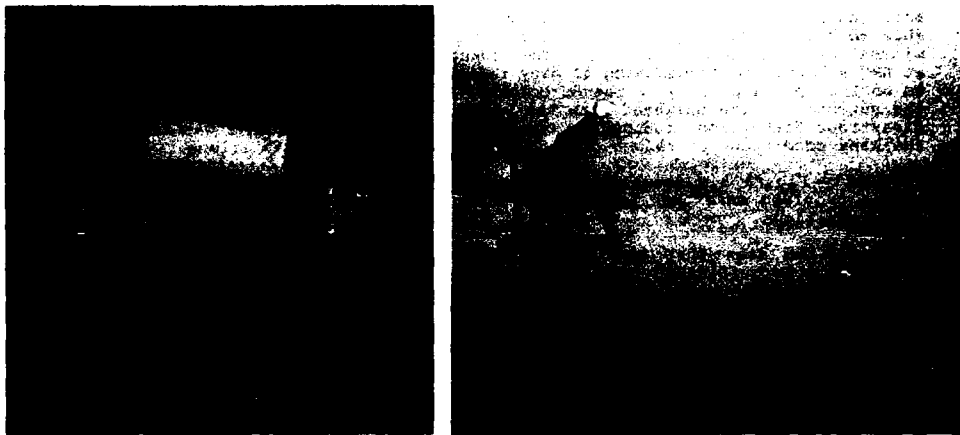


Figure 5.5 - Challenger Nose Section and Antenna Mockup

5.2 FULL-SCALE MOCKUP AND MEASUREMENTS

A wooden aluminum-covered mockup of the first two frames of the aircraft nose section was constructed as shown in Figure 5.5. The actual weather radar antenna was not available, hence a mockup of its support and plate array was constructed as well. The top half of mockup was extended sufficiently to allow the mounting of the transmitting horn to emulate the top-mounted fuselage antenna location. The off-centre hole in the upper portion of the bulkhead is for the mounting of the second jammer antenna. The structure on the other side of the scanner is an absorbing baffle whose function is discussed below. The second photo shows the mockup on the roof-top range with the flooding horn antenna opposite.

Only highlights of the measurement results will be presented in this paper. Of particular interest is the elevation pattern of the top-mounted antenna. Figure 5.6 shows the elevation plane patterns of the low and high gain antennas in this position. Note the normally smooth elevation pattern is characteristically distorted by the fuselage reflections from the sloping section. Coverage would be poorer for a flatter fuselage section.

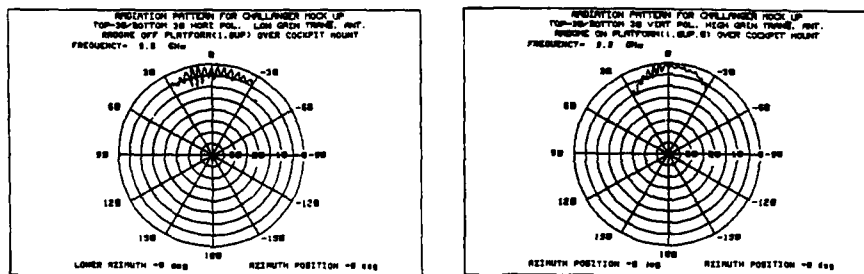
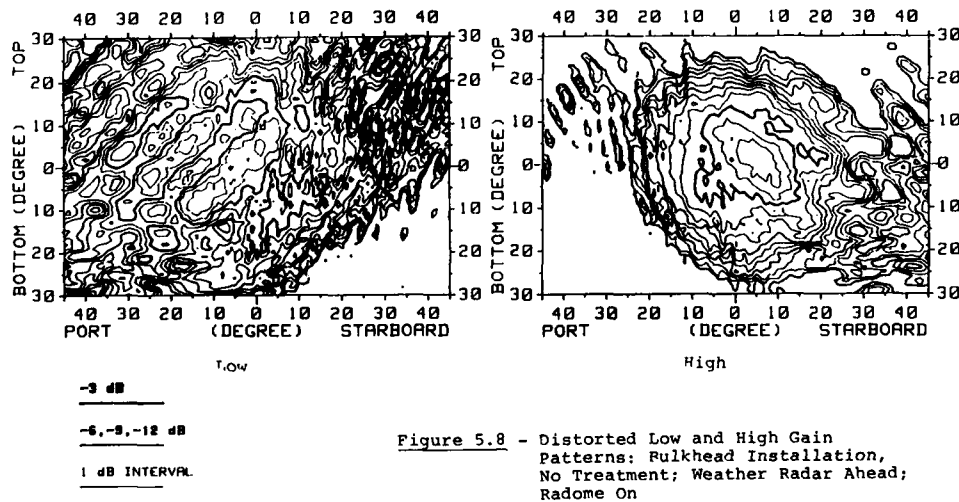
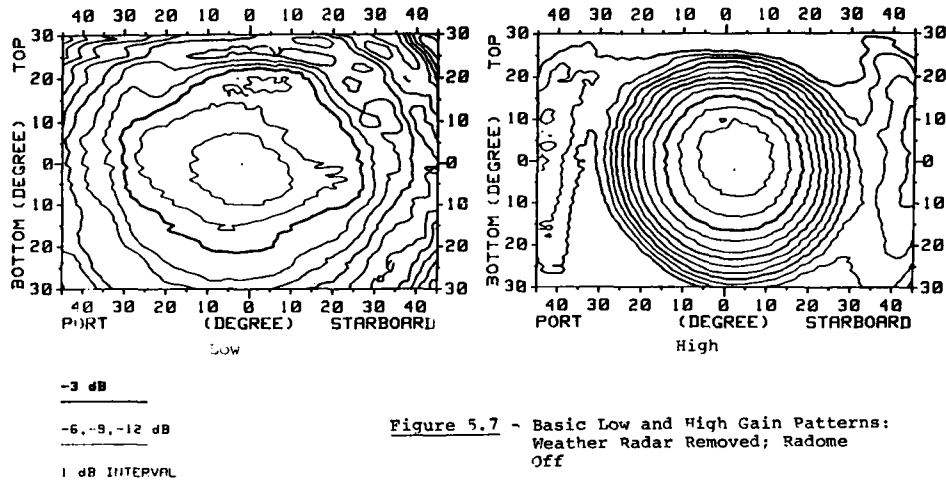


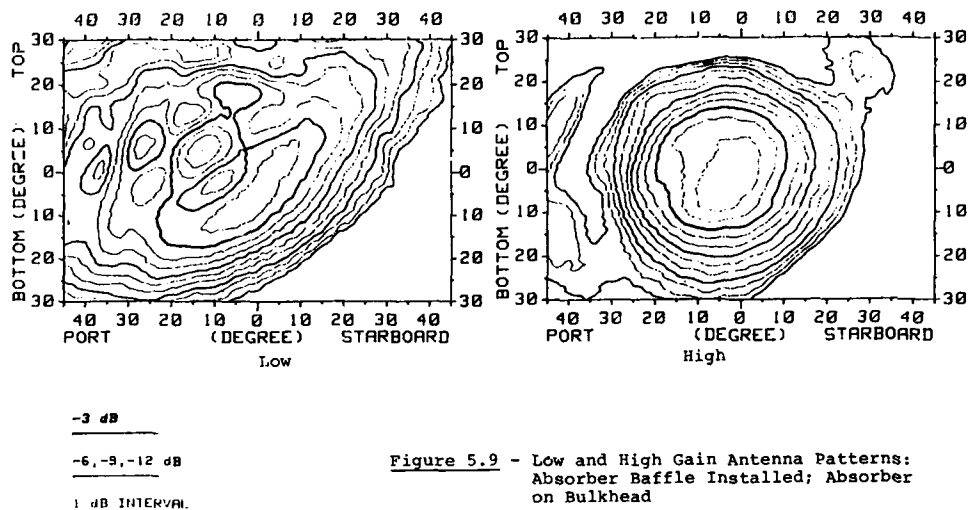
Figure 5.6 - Elevation Plane Patterns: Low and High Gain Antennas

Transmit to receive antenna coupling within the radome can be controlled by a baffling arrangement. If the baffle were to be of absorbing material, it might be able to control the expected radiation pattern distortions sufficiently to have a minimal level of forward coverage. The computer-based HP8510 Network Analyzer system at NRC allowed the programming of many radiation pattern cuts. These are presented as contour plots (usually in pseudo-colour). Figure 5.7 shows the smooth patterns of the antennas on the bulkhead with the weather radar removed. Figure 5.8 shows the disastrous distortion produced by the weather radar system. This was tested under the many conditions of radome on, and scanner in various orientations.

Finally Figure 5.9 shows the results with the absorbing baffle in place. The baffle provided more than the minimum level of antenna decoupling for all weather radar positions.

It can be seen that this sequence exposed the necessary radiation pattern information in sufficient detail for the operational effects to be identified.





6. CONCLUDING DISCUSSION

The sequence of examples which had been discussed was intended to expose the possibilities and limitations of the computational modelling techniques in their application to very practical cases. In cases such as these the required time response must be kept in mind. To this end rapidity of model generation and a flexible system of analysis of the results is important.

It can be seen that the best combination is a parallel measurement and computational modelling program. Also exposed were the problems of source representation for practical antennas and the level of complexity in the geometry which must be addressed. It is important to use every measurement experience to enrich the measurement data base which could be made available for the validation of computational electromagnetic modelling.

It is important to conclude that as in the helicopter case, computational models can provide meaningful radiation pattern information which can be used for the analysis of operational scenarios, and which cannot be obtained in any other way.

7. ACKNOWLEDGMENTS

The work described in this paper was primarily supported by the Defence Research Establishment Ottawa (D.R.E.O.) under the sponsorship of Drs. P. Bhartia, David Liang and Satish Kashyap and of Messrs. R. Kirk and Paul Campagna of DND. The helicopter and Challenger measurement information was supplied by John Hazell, Dr. S. Mishra and Randolph Balaberda of the Electromagnetic Engineering Section of the National Research Council. The ship modelling work was done for Mr. Ron Roe of DND. Dr. E. Cerny, Dr. C. W. Trueman, David Gaudine, Colin Larose, Lan Nguyen, Sergio Prestipino, Duy Nguyen, Vito Salvagio, Victor Phan, Benita Rozensweig and Michael Wong and Khiet Le of the EMC Laboratory were enthusiastic contributors to the research activity.

It is always important to acknowledge the developers of the NEC computer code and the Basic Scattering Code for providing the practical basis for this type of modelling effort.

The work was also partially supported by an operating grant from the National Sciences and Engineering Research Council.

REFERENCES

- [1] Miller, E. K. "A Selective Survey of Computational Electromagnetics for Antenna Applications", AGARD Lecture Series No. 165, October 1989.
- [2] Burke, G. J., "Recent Advances to NEC: Application and Validation", AGARD Lecture Series No. 165, October 1989.
- [3] Molinet, F. A., "GTD/UTD: Part 1. Brief History of Successive Development of Theory, and Recent Advances", AGARD Lecture Series No. 165, October 1989.
- [4] Audone, B., "Compact Ranges for Antennas and RCS Measurements", AGARD Lecture Series No. 165, October 1989.
- [5] Bothe, H. and Macdonald, D., "Determination of Antennae Patterns and Radar Reflection Characteristics of Aircraft", AGARDograph No. 300, AGARD Flight Test Series Vol. 4, May 1986.
- [6] Ibid, p.13, Figure 1.19
- [7] Ibid, p. 14, Figure 1.21
- [8] Kubina, S. J., "Numerical Modelling Methods for Predictiong Antenna Performance on Aircraft", AGARD Lecture Series 131, The Performance of Antennas in Their Operational Environment, October 1983.
- [9] Kubina, S. J. and Trueman, C.W., "Selected Applications of Method of Moments Modelling", NATO ASI on Modelling and Measurements, Il Ciocco, Italy, August 1987.
- [10] Wu, T.T. and King, R.W.P., "The Tapered Antenna and its Application to the Junction Problem for Thin Wires", IEEE Trans. on Antennas and Propagation, Vol. AP-24, No. 1, pp. 42-45, January, 1976.
- [11] Bhartia, P., Kubina, S.J., Cerny, E., and Gaudine, D., "Generation of Computer Models for the EMC Analysis and Design of Large Systems". Proceedings of the 1984 International Symposium on Electromagnetic Compatibility, Tokyo, Japan, October 1984.
- [12] Trueman, C.W. and Kubina, S.J., "AM Re-Radiation Project - Final Report", Technical Note EMC-80-03, Concordia University, EMC Laboratory, March 1980.
- [13] Halpern, Ben and Mittra, Raj, "On the Validity of Terminal Properties of Thin-Wire Antennas Computed Using the Numerical Electromagnetics Code(NEC)", Program Abstracts 1986 National Radio Science Meeting, Philadelphia, Pa.
- [14] Popovic, B.D., Dragovic, M.B., and Djordjevic, A.R., "Analysis and Synthesis of Wire Antennas", Research Studies Press (Wiley), 1982.
- [15] The Applied Computational Electromagnetics Society, Dr. Richard K. Adler, Secretary, Naval Postgraduate School, Code 62AB, Monterey, CA, 93943 USA.
- [16] Kubina, S.J., "Helicopters, HF Antennas: Models, Measurements and All That", Proceedings of the ACES Conference, Monterey CA, Apr. 1987.
- [17] Law, Preston, E., "Shipborne Antennas", Artech House Inc. 610 Washington St. Dedham Ma. 02026.
- [18] Marhefka, R.J. and Burnside, W.D., "Numerical Electromagnetic Code - Basic Scattering Code, NEC-BSC (Version 2), Part I - User's Manual", Ohio State REport 712242-14, December 1982.
- [19] Larose, Colin, "Advances in Wire-Grid Modelling of Antennas and Auxiliary Computer Graphics Systems with the CP-140 Case Study", M. Eng. Thesis, Dept. of Electrical Engineering, Concordia University, July 1986.
- [20] Trueman, C. W., Pavlasek, T. J., and Kubina, S. J., "Parametric and Synoptic Charts for Determining the Patterns of an Elemental Dipole with a Scattering Mirror or Cylinder", IEEE Trans. on EMC, Vol. EMC-19, No.4, Nov. 1977.
- [21] Kubina, S. J. and Larose, C., "A NEC Topside Antenna Case Study with DIDEK and SPECTRUM: Model Generation and Current Display Codes", Proceedings of the ACES Symposium, Monterey CA, March 1986.
- [22] Westman, H. P. (Ed.), "TACAN Symposium Issue", Electrical Communication, ITT Corp. Vol. 33 No.1, March 1956.
- [23] Kubina, S. J. and Drolet, M., "Modern Aids to Effective EMC Testing", Proceedings of the 1989 Flight Test Symposium of the Canadian Aeronautics and Space Institute (CASI), AETE, CFB Cold Lake, Alberta, 8 March 1989.

HF Wire Antennas over Real Ground: Computer Simulation and Measurement

John S. Belrose and G.M. Royer
Communications Research Centre
Ottawa ON K2H 8S2

L.E. Petrie
Petrie Telecommunications
Ottawa ON K2J 1G4

SUMMARY

This lecture is based on the considerable combined experience of the authors with numerical and experimental modelling, with fabricating, with measuring performance of and with using practical HF antennas for the low HF band (2-8 MHz). The antenna poses a particularly difficult problem for this frequency band, particularly if one wants a broad-band antenna for frequency agile systems, because of the four-to-one change in wavelength over the band. A further complication is that as the frequency is reduced so is the electrical height of the antenna above the ground, and since the ground is an imperfect conductor, the finite conductivity of the ground introduces loss of gain, and detuning.

We have been particularly concerned with antennas at low (electrical) heights, even with antennas very near to or actually lying on the ground. Antennas lying on or near the ground are of particular interest for military tactical communicators, since low profile antenna systems that require no support mast or masts are operationally and logistically an advantage for use in a tactical environment. HF radio operators in the Canadian north have frequently reported success with using HF dipoles lying on the ground. The reason that antennas perform at all under such circumstances is because of the very low ground conductivities found in the Canadian north. In this lecture we have shown (for the first time in the authors experience) what gain to expect under such circumstances.

We have in keeping with the theme of the lecture series particularly emphasised the comparison between computer simulation with measurement, and with the application of computers for both numerical modelling, measurement, and data analysis and presentation.

1. Introduction

This lecture is concerned with the performance of practical HF antennas in their operational environments, with particular emphasis on wire antennas at low HF frequencies (2-8 MHz). In the 1950s and 1960s an important role for HF communications was to provide long distance fixed links for global communications. Since the advent of satellite communications HF radio is increasingly being used to provide short and medium distance communications, especially between mobile or transportable stations. The propagation characteristics for these shorter links require the use of correctly chosen frequencies, and antennas which radiate efficiently at the required angles of elevation, viz antennas which are suitable for near vertical incidence sky-wave (NVIS) links.

For military communications this band of frequencies is used for land, air and marine mobile, and for tactical communications. Vertically polarized antennas are required for ground wave paths, to extend the range to distances beyond the radio horizon, to distances beyond those reached by VHF; and horizontally polarized antennas are required for high angle sky-wave links to complement VHF which has a shorter range and is strongly affected by terrain features. For civilian applications HF short to medium range links, to distances of several hundred kilometers, are used for rural and remote area radio communications. For such applications the performance of the antenna can impose important limitations on the quality and reliability of radio communications. Employing high transmitter powers to overcome the poor radiation efficiency of an antenna is not a recommended route to take, since this only creates other problems, and there is an associated increase in cost; and after all the antenna has to used to receive. If you cannot hear the station you wish to communicate with you will not be able to communicate no matter what your transmit power. For mobile communications, the radiation efficiency of the antenna becomes an important parameter, since an 8-to 12-foot whip, which is typically employed for land mobile, or a 35-foot whip for marine mobile communications, is an antenna that is electrically short, particularly at low HF frequencies. An important limitation, however, in the present context, is the overhead radiation pattern null, which makes this antenna not very suitable for high angle sky-wave links. The overhead null problem can be overcome through the use of a horizontal dipole providing its height above the ground in wavelengths (h/λ) is less than, and does not approach $h/\lambda=0.5$.

This antenna however has the following problems associated with it:

- For the lower frequencies in the HF band practical dipole heights may make h/λ sufficiently small such that losses, in the non-perfectly conducting ground, make the antenna unacceptably inefficient;
- Unless an antenna system tuning unit (ASTU) is employed, a dipole has, due to it changing input impedance as a function of frequency, a bandwidth of only about 5-percent. Note that the desired bandwidth previously mentioned covers a four to one frequency ratio; and
- A horizontal dipole of half length $< 5/8\lambda$, whose height above the ground is sufficiently less than $\lambda/2$, has a main lobe whose maximum is in the vertical direction. As the frequency is increased for a dipole of a given dimension at a fixed height above the ground, the above antenna will develop a multi-lobed pattern. For example, its pattern in the vertical direction will correspond to: a) the peak of a minor lobe when its half length is 0.75λ ; b) a pattern null when its half length is 1λ ; and iii) there is an overhead pattern null when the dipole height is 0.5λ and 1λ .

This lecture is concerned with the operational performance of practical HF antennas, as judged by measurement and numerical modelling, using the computer program NEC-3 [see Burke, 1980]. We will address the subject of using the antenna itself to estimate the conductivity of the ground beneath it, which affects performance, and particularly with a comparison between numerical modelling and measurement in full scale. Numerical modelling is an important new tool for antenna design/antenna optimization, since a parameter study is much more easily done by numerical modelling than by constructing antennas and measuring performance in the real world.

2. The effect of real ground on antenna performance

2.1 Electrical properties of the Ground

The electrical characteristics of the surface of the earth are determined by the nature of the soil, by its general geological structure, by the moisture content, and by the frequency. Since this is an overview paper, it needs to be noted that various authors are not consistent as to the conductivity (σ in S/m) and the relative dielectric constant (ϵ_r) values that are used for HF for so-called good, average and poor ground. Some of the antenna performance parameters (impedance, pattern, efficiency, gain) given in this lecture have been calculated using CCIR recommendations for ground conductivity [Recommendation 527, CCIR XVth Plenary Assembly p. 73-75, Dubrovnik, 1986]; others have employed values proposed by G. Hagn, SRI International, e.g. those of Breakall [private communications, 1986]. See Table below:

	CCIR Recommendation		G. Hagn, SRI Int'l	
	σ (S/m)	ϵ_r	σ (S/m)	ϵ_r
Wet Ground	10^{-2}	30	5×10^{-2}	42
Average Ground			3×10^{-3}	13
Medium Dry Ground	10^{-3}	15		
Very Dry Ground	10^{-4}	3	1.5×10^{-4}	4

2.2 Effect on Radiation Pattern

The lossy ground beneath antennas affects performance. If ground conductivity is poor it is in general better, depending on pattern requirements, such as desired launch angle and height for practical antennas, to employ horizontal polarization rather than vertical polarization; since horizontally polarized radio waves are less affected by the finite conductivity of the ground beneath and in front of the antenna.

For perfectly conducting ground the voltage reflection coefficient for **horizontally polarized** waves equals -1 for all angles of incidence. A horizontal current element thus has a reflected image in which the current flows in the opposite sense and ground reinforcement occurs at angles of elevation for which the path difference between direct and reflected waves is one-half a wavelength. The dominant

parameter determining the pattern is the mean height of the active part of the antenna above ground. If the ground is not perfectly conducting, the reflection coefficient at small launch angles has a smaller value, but the phase remains near to 180° , so the pattern is scarcely changed. The actual reflection coefficient may lead to a loss of gain, usually less than 3 dB, but even here the directive gain remains 3 dB higher than it would if the ground reinforcement were absent. The launch angle depends on the height of the antenna and the antenna type used. For a horizontal dipole at heights of 0.2, 0.4, 0.5, 1, 2, and 3-wavelengths the launch angle changes from 90° to 40° , 30° , 14° , 7° , to 4.7° respectively (c.f. the vertical-plane patterns in Fig. 1).

For vertical polarization the effects are more complex. For highly conducting ground, the reflection coefficient is approximately 1. A vertical current element above the ground has a reflected image in which the current flows in the same sense, so ground reinforcement occurs at low launch angles. When the conductivity is finite, the phase of the complex reflection coefficient changes by 180° at small launch angles. Over sea this happens at launch angles of a fraction of a degree above the horizon, and the formation of the low angle maximum is scarcely affected. Over land the phase change occurs at launch angles around 15° , and the reflection coefficient falls to a low value in the region of the change. The angle which this occurs is called the Brewster angle. Antenna performance is seriously affected, particularly as the minimum reflection coefficient occurs at angles which are comparable with the most probable angle of arrival of HF sky-waves propagated over long distance paths. For propagation paths >2000 km the optimum launch angle are $<15^\circ$. A further consideration is that a vertical monopole must be fed against a radial ground system or against an elevated radial ground system, or counter-poise, and ground conductivity effects the design of ground systems.

Thus, while an ideal vertical monopole/dipole in theory offers the possibility of a low radiation angle, practically a horizontal dipole may be the preferred polarization, because of problems associated with ground conductivity, unless the launch direction is over sea water. An important consideration for the latter antenna, however, is that the height required for optimum launch angle for a long communications path may not be practical. As noted above, the antenna height for a dipole must be $> \lambda$ if the launch angles desired are $<14^\circ$; and this is certainly not practical when low HF frequencies are employed. For example, a 30 metre mast is a practical height for a support mast. For a frequency of 10 MHz this mast height corresponds to one wavelength, and so the launch angle for a horizontal dipole at this height would be 14° . However for a frequency of 4 MHz it would be necessary for the same launch angle to employ a mast height of 75 metre, which is impractical. Thus, in spite of the fact that the performance of a vertical monopole may be poor at low launch angles, it may out-perform a practical horizontal dipole. The diagrams in Fig. 2 illustrate the points discussed above. Here the gain for horizontal dipoles at 60 feet (18m) and 100 feet (36m) for two frequencies, 3.8 and 7.2 MHz, are compared with a quarter wave monopole for good and poor ground conductivity. The various curves for the monopole are for different numbers n of radial wires comprising the radial wire ground system ($n=30, 60$ and 120), see note for Fig. 2.

2.3 Effect on Impedance

A practical "half-wave dipole" refers to a resonant-length dipole for which the input reactance (X_a) is zero. Because of end effects the resonant length is always shorter than a half-wavelength by an amount depending on the antenna thickness, and the radiation resistance is somewhat less than the theoretical value of 73.1 ohms, which corresponds to that of an infinitely thin dipole in free space. It is interesting to note that NEC-3 gives a resonant input resistance value for a 4-MHz wire dipole in free space of 67.0 ohms.

In addition to the above effects, an actual dipole is always located above the ground or other supporting and reflecting surface, so that the theoretical free-space conditions do not apply. It has long been known that the presence of a finitely conducting ground and the dipole's supporting structure can significantly effect the input impedance. A selection of some of the early work is reproduced in Fig. 3. Curves 1, 2 and 3 are for perfect ground. They are respectively, the early theory of Friis et. al... [1934]; the more rigorous theory of Miller et. al... [1972]; and the experimental results of Proctor [1950]. The data points (open circles and triangles) are the measurements of Friis et.al. at wavelengths from 8-27m for vertical and horizontal dipoles over real earth. The dashed curve 4 is fitted to their data (the open triangles) at low heights. Curve 5 was measured by Hagn and Baker [1972] at 6 MHz (50m wavelength) for a horizontal dipole over good ground ($\sigma=25\text{mS/m}$). Bhattacharyya [1963] has given numerical results that are in qualitative agreement with these data.

In general the input impedance of a dipole over real ground has a reactive component as well as resistance, and the magnitudes of both resistance and reactance oscillate about the free space values, with change in the height of the antenna above the ground. While this effect has also long been known, it is only recently that numerical modelling has achieved a sufficient rigour that one can have confidence in the calculated result. It is interesting to know the magnitude of the effect. In Fig. 4 we show the results for such a theoretical calculation (using the computer program NEC-3) for a real physical dipole above good ground ($\sigma=10^{-2}$ S/m). According to NEC-3 this dipole was not resonant even in free space. This is because the dipole's half length was computed using the equation $0.95 \lambda_0/4$, where λ_0 is the free space wavelength. This is the usual equation used by antenna engineers for practical HF dipoles; whereas for resonance, according to NEC-3, the equation $0.975 \lambda_0/4$ should have been used. Because of this problem with resonant length the reactance scale of Fig. 4 was normalized so that zero corresponds to the theoretical value of the antenna's reactance in free space. In effect this meant that we had to add 42.36 ohms to all reactance values predicted by NEC-3. Thus, without having to recalculate, the NEC-3 results throw a new light on the dependence of input impedance on the height of the dipole. The graph without this normalization is not shown. In particular it is interesting to note that the dipole is first resonant ($X_a=0$) when its height is about $\lambda/3$, which is the same height where its input resistance is a maximum. In fact the reactance is zero wherever the resistance is a minimum or a maximum. Intuitively this is not surprising, but the height which this occurs is not simply related to in-phase or out-of-phase relationships between the direct and ground-reflected waves on simple theory. The fact that the antenna is reactive, except at certain heights and in free space if its length is not right explains why HF antennas, if one is a perfectionist and wants the antenna to be exactly resonant, has to be "pruned" to length on site suspended from the tower(s) that support it.

In Fig. 5 we show measured impedance ($Z_a=R_a \pm jX_a$) values for this dipole versus height (in wavelengths). It was a 4 MHz dipole (fabricated of Number 12 antenna wire) over a flat mowed grassy site. The purpose of carrying out the measurements graphed in this figure was to compare measurement with theory, since if the agreement was satisfactory, then it was surmised the impedance of a dipole at low height could be used to estimate the ground conductivity beneath it. The ground conductivity at the measurement site was estimated to be 10 mS/m, based on previous experience. There is a small but systematic difference between the summer and winter values. The seasonal change is just detectable, in spite of the fact that the conditions were quite different. In summer the measurements were made when the ground was quite damp. The winter measurements were made when the ground was frozen and covered with snow.

The dependence of dipole impedance on dipole height (h/λ) and ground conductivity according to theory is given in Fig. 6 (see comments on that figure), for three quite different grounds: very poor, average and good ground. The heavy curves are resistance, R_a ; the light curves are reactance, X_a . Clearly the resistance values have a marked dependence on ground conductivity. There is only a small change in reactance with ground conductivity. The data points not connected by a curve are the measured values. Open circles are resistance, the solid circles are reactance values (the summer values of Fig. 5). The measured resistance values are in good agreement with the theoretical curve for $\sigma=10^{-2}$ S/m, at least at higher heights. A slightly greater value for ground conductivity might yield a better agreement between theory and experiment at lower heights. The measured reactance values are also in approximate agreement with theory, at least the variation of reactance with h/λ is similar. It should be noted, however that in the figure we are not comparing dipoles having the same physical length. We are comparing resonant dipoles. The theoretical curves are for a dipole that is 2.5-percent shorter than a half wavelength (which is the resonant length for a dipole in free space according to NEC-3); the experimental dipole was 5-percent shorter than a half wavelength (which is the resonant length for a practical HF dipole at heights typically employed). We have discussed this difference above (see Fig.4).

2.4 Effect on Gain

As noted at the outset, transportable HF antennas are frequently operated at relatively low heights above the ground; and often a drooping dipole is employed rather than a horizontal dipole, since a dipole with drooping ends requires only one mast for support rather than two. In this section we compare gains measured for propagation over a NVIS path, between drooping and horizontal dipoles, for three frequencies and for different antenna heights. The measured gains were relative wrt resonant horizontal dipoles at 30-feet (9.1m). The results graphed, however, are gain in dBi. This has been determined by theoretically calculating the gain for the reference dipoles, for an assumed ground conductivity of 10^{-3} S/m (sandy soil). The measured gains graphed in Fig. 7 show that height above the ground is a more

important parameter than is frequency. The measured gains graphed in Fig. 8 (here relative gains are plotted) show that the drooping dipole is just as good a radiator as a horizontal dipole, provided that the ends of the dipole are kept well off the ground. The surprising result shown by data points 1 and 2 are for a drooping dipole at 12-feet (height of peak), with ends at 5-feet and 2-feet respectively. These measurements were made for a frequency of 4 MHz, however similar results were obtained at the two other frequencies.

Hagn and Barker [1970] conducted a careful series of measurements of the gain toward the zenith of a half wave dipole as a function of antenna height, at several sites. Their results are more conducive, then ours, for comparison with theory, because they measured the ground conductivity at the test sites by a two probe rf experiment. Some of their measurements were made in a clearing, some in forested sites, some after a rain, since the study was particularly concerned with the effects of the jungle. We have selected one set of measurements made in a clearing, at Chumphon, Thailand. The ground conductivity was very high. This particular set of data was chosen for another reason, measurements were made to very low heights (1 foot or 0.3m). The measurement frequency was 6 MHz. In Fig. 9 we compare their measured gains (open circles) with our calculated gains (according to NEC-3). The curves are the theoretical gains for two ground conductivities measured by the authors on two different occasions. The conductivity parameters $\sigma = 58$ mS/m, $\epsilon_r = 53$ give the best agreement with measurement, however there is a systematic departure between theory and experiment at low heights (below about 0.025λ). The experimental gains are less than the gains calculated for either conductivity. It is worth noting, however, that at very low heights the antenna is quite detuned, and it is not known whether the authors used an antenna system tuning unit.

The ground conductivity at a test site can be estimated by measuring the difference in gain for a dipole at two quite different heights. The vertical incidence gain for a dipole at low heights (say $h/\lambda = 0.006$) decreases as ground conductivity increases, whereas the gain at a higher height (say $h/\lambda = 0.1$) increases as conductivity increases (see Fig. 10). This difference is intuitively in the right sense. Thus the difference in gain between a dipole at a low and a high height increases differentially as ground conductivity increases. For example, for a 4 MHz dipole at 0.5 and 7m, this difference in gain amounts (according to NEC-3) to 4.1, 10.9 and 14.9 dB for ground conductivities of 0.1, 1 and 10 mS/m respectively. Thus when carrying out an experiment to measure the NVIS gain for an antenna under test, the effective ground conductivity at the test site can be estimated by this method.

2.5 Bandwidth

The operational bandwidth of a transmission system depends on the type of antenna employed, to some extent on the ground conductivity beneath the antenna, and on the ability of the transmitter to deliver power to a reactive load. Some antenna systems have a wider bandwidth than others. In this overview section we have been discussing basic antennas, viz. dipole antennas. It was mentioned at the outset that the bandwidth of a wire dipole is about 5-percent. But it was not mentioned how bandwidth was defined. Antenna manufacturers normally specify antenna bandwidth in terms of the voltage-standing-wave-ratio (VSWR) being less than a specified value, say $VSWR < 1.5$ or $2:1$. But what does this mean? Modern solid state transmitters usually have a VSWR detector, and when $VSWR > 1.5:1$ the transmitter output power is reduced, typically the output power is reduced by about one-half (3dB below nominal power) when the $VSWR = 2:1$. The theoretical bandwidth (according to NEC-3) of an HF wire dipole antenna in free space defined as VSWR being < 1.5 or $2:1$ is 3.3 and 5.9-percent respectively (see Fig. 11). But this is not the antenna's bandwidth, it is an operational bandwidth defined according to the performance of a transmitter. If one thinks of an antenna in terms of an equivalent L-C-R circuit, the half-power bandwidth would be determined from the difference between the two frequencies where the impedance phase angle $\theta \pm 45^\circ$. On this definition $R = \pm jX$, and the bandwidth is 8.6-percent. If one defines the antenna bandwidth in terms of its half-power bandwidth, then when the $VSWR = 5.8$, one-half the power from the transmitter incident on the antenna terminals, enters the antenna. For $VSWR < 5.8$ the antenna's bandwidth is 16.8-percent [Royer, 1989]. Clearly this latter bandwidth is only of academic interest, since practical transmitters will deliver negligible power to a reactive load having such a high VSWR. If an antenna system tuning unit (ASTU) is employed the system bandwidth would be greater, defined by the tuning range of the ASTU and the acceptable loss due to VSWR on the transmission line feeding the antenna.

The off-resonant VSWR for an actual antenna over real earth will be less than that calculated for a lossless antenna in free space (and hence its bandwidth will be correspondingly greater), for two reasons: 1) loss is introduced by the finite conductivity of the ground beneath the antenna (the antenna

resistance R_a is equal to the radiation resistance R_r plus a ground loss resistance R_g ; and 2) transmission line loss. The measured VSWR for author Belrose's 4 MHz dipole at 50 feet (15m), albeit a drooping dipole, is plotted on the corresponding graph of Fig. 10 (the dashed line). The difference cannot be explained by transmission line loss alone. For the 60 feet (18m) of RG8U coaxial cable used, the transmission line loss (at 4MHz) is about 0.24 dB, and for this loss, for a VSWR at the transmitter equal to 4:1, the VSWR at the load (the antenna) would be about 4.5:1. The graph predicts a VSWR of 6:1. The measured bandwidth for VSWR<2:1 is about 7.5-percent. Clearly we should have been comparing what NEC-3 would have predicted for the actual real situation.

3. Practical Wire Antennas of Military Interest

It was mentioned at the outset that an important application for HF is for near vertical incidence sky-wave (NVIS) links, for short to medium range radio communications. The communications reliability and quality of mobile and transportable HF communication terminals is commonly determined by the poor performance of the antennas employed. Improved performance requires transportable antennas, which ideally, must exhibit seemingly contradictory characteristics. The antenna must be extremely light weight for ease of conveyance, yet rugged enough to withstand the rigors of transport and installation. It must be simple in design and essentially pre-assembled to allow rapid installation by a small crew, yet it must use materials and connectors which allow the antenna to unpack/and pack easily, and it must perform reliably. The disassembled antenna must be compact enough to fit in a small, easily transportable container, yet the erected structure must have a large radiating aperture if high (even acceptable) power gain is to be realized.

From an electrical characteristic point of view (gain-bandwidth response), one of the inherently broad-band antennas would best meet the requirements for frequency agile HF communications. However because of their complexity and erected physical size, they are not normally considered by military and civilian communicators as meeting the requirements for mobile, portable or transportable communications. Notwithstanding, considerable ingenuity and design effort has been devoted to developing (so called) compact, easily erected versions of such antenna systems. Examples are the TCI Model 537 transportable log-periodic dipole array; and the Granger Associates (now Andrews Antenna Company) 'Spira-cone' Model 3002 antenna. In the view of the authors, however, these antennas are not considered to be easily erected or compact. They are considered to be more suited for fixed location operation. The present study is therefore confined to the following simpler structures: 1) tuned dipoles and loops; and 2) resistance loaded traveling wave type of antennas. The latter are broad-banded, and do not need to be tuned, but suffer from reduced radiation efficiency due to power absorbed in the loading resistors.

3.1 Tuned Antennas

Since vertical whips do not illuminate high elevation angles, NVIS antennas must use horizontal elements or loops.

The most satisfactory antenna is perhaps an ordinary half-wave dipole mounted about 0.35-wavelength above the ground, and bent in the vertical plane to avoid nulls off the ends. However wire dipoles are relatively narrow band antennas (without tuning the operational bandwidth is about 5-percent), whereas the antenna must be useable over at least a 4:1 frequency band (say 2-8 Mhz). Current demands for transportable antennas having wideband performance require design which must make significant performance compromises: 1) large antennas radiate efficiently at the lower frequencies, but usually have multi-lobed radiation patterns at the higher frequencies; 2) small antennas have well-behaved radiation patterns, but their efficiency and bandwidth fall rapidly as the frequency is reduced; 3) antennas mounted above the ground have elevation patterns which vary with frequency; if the height is reduced too far, efficiency falls due to losses in the ground; and 4) modern microprocessor controlled antenna system tuning units (ASTUs) are available, but these are designed for unbalanced antenna systems, where dipoles are a balanced antenna; and the operational parameters for ASTUs are in general such that they cannot cope with the large range of impedance values near the anti-resonant frequency.

The severity of some of these effects is why communicators and antenna engineers are still looking for an optimum antenna design.

3.1.1 Compact Loop Antennas

Electrically small loops (perimeters as small as 0.04 to 0.3-wavelength) are characterized by a very small radiation resistance, and therefore such loops must be fabricated from large diameter tubing (2-10 cm diameter tubing is typically employed) to keep losses small. Since such loops are inductive they can be readily tuned by a capacitor, and power coupled into them by means of a small coupling loop (see Fig. 12e). Patterson [1967] is attributed to be the first to suggest using compact loops for tactical communications. He used a capacitor-match arrangement, which is an efficient but more complicated method of tuning and matching than mentioned above.

The far-field signal radiated by a loop antenna (whether in free space or above a ground plane) is polarized parallel to the plane of the loop. The gain pattern, for a loop which is mounted such that the loop's plane is normal to that for the ground, is essentially omnidirectional for angles near the vertical. The pattern is the typical figure-of-eight for angles near the horizon. If the plane of the loop is horizontal, the far-field radiated signal will be horizontally polarized and the gain pattern azimuthally omnidirectional. The vertical pattern will depend, as for any horizontally polarized antenna, on the antenna's height above the ground plane.

If the operating frequency band is restricted to say 2:1, and the physical size of the loop is not too small (say perimeter $0.12-0.25\lambda$) acceptable gains can be achieved (4-6 dBi for a vertical loop at low height over perfect ground); but the gain falls off rapidly as frequency is reduced, i.e. perimeter/ λ becomes small, and the bandwidth becomes small. Since the radiation and loss resistances are low the current is high. A vacuum variable capacitor must be used, and the loop must be tuned for exact resonance. For example, the Racal compact loop, which is one of the antennas that has been tested by the authors (see below), has the following operational characteristics at a frequency of 4 MHz: 5.4 and 8.5 milliohms radiation and loss resistances respectively; gain over perfect ground 0.7 dBi (measured NVIS gain over sandy soil -2.7 dBi); antenna bandwidth 0.7 kHz; and the loop current for a transmitter power of 1-kW would be 267 amperes. The voltage across the capacitor would be 21 kV. The operational bandwidth will be greater than the antenna bandwidth because the output impedance of the transmitter is in effect across the input terminals of the antenna (for matched conditions the bandwidth is doubled); and because of ground loss resistance. Ground loss resistance is in effect an additional dead loss resistance in series with the radiation and copper loss resistances. For small loop antennas that are operated over a finitely conducting ground the ground loss resistance is comparable to the other resistances in the circuit. Notwithstanding, the BW is small, and it becomes vanishingly small for low HF frequencies, particularly for small loops that are operated over sea water.

3.1.2 Dipoles

A microprocessor controlled automatic antenna system tuning unit (ASTU) can be used to extend the working bandwidth of a dipole. A dipole 1.25-wavelength long at 10 MHz is still 0.25-wavelength long at 2 MHz, and will operate efficiently over the range if provided with a well-designed ASTU. The elevation pattern of the dipole becomes narrower and its gain increases as the frequency increases, but this is partly offset by the effect of ground reflection. A mounting height chosen to optimize the pattern at 10 MHz (say 0.35λ) is very low at 2 MHz (0.07λ) and large ground losses result.

Royer [1987a,1987b] has made a detailed numerical modelling study of a switched drooping dipole with tuned feeders (see Fig. 12c); and a switched vee-dipole (see Fig. 12d), which requires no feeders for a tuning unit at ground level. We show in Fig. 13 a comparison between Royer's calculated values for impedance, for a switch drooping dipole with tuned feeders and measured values, at various frequencies (not shown) over the band 2-10 MHz (the lowest, highest and mean frequency for each of the five sub-bands. There is a clear correlation between calculation and measurement. While there is not a one-to-one agreement we should not expect there to be, since the calculations were for a ground conductivity of 8 mS/m ($\epsilon_r=15$), whereas the measurements were for an antenna over sandy soil ($\sigma=1$ mS/m). The data scatter is in part due to measurement inaccuracy. Since a balanced-impedance-meter was not available the balanced antenna system impedance was measured in a tedious somewhat approximate way. The antenna was tuned-and-matched for an input impedance of 50Ω using a balanced manually tuned ASTU. The ASTU was then removed from the circuit. Its input was terminated in 50Ω and the impedance seen looking into its normal output terminals was measured. This impedance is the conjugate impedance of the antenna system. The accuracy of measurement depends on the accuracy of tuning the ASTU.

The undesirable addition of an ASTU; and the requirement to switch the dipole (in effect change its length as frequency is changed in accordance with 4- or 5-sub-bands) complicates the antenna design (increasing the cost and complexity of the system). Presently R&D activity is toward development of a microprocessor controlled balanced ASTU, and toward devising a basic fan type dipole to reduce the large impedance maximum at the anti-resonant frequency.

A variety of fan dipoles has been used for NVIS applications. These have taken the form of flat bow-ties or three dimensional pseudo-conical arrangements. The problem with this type of antenna is that it is somewhat complicated to untangle and erect, and while well designed antenna systems are available, they are expensive. These antennas suffer major radiation pattern breakup at higher frequencies, a factor which, because their VSWR bandwidth is large, is often overlooked.

3.2 Travelling wave type Antennas

The standard dipole's nearly sinusoidal current distribution is a standing wave produced by two nearly equal amplitude, travelling waves moving in opposite directions along the antenna's wires (the arms of a dipole). We will refer to the above waves as primary and secondary travelling waves where: a) the primary travelling wave originates at the dipole's feed and moves toward the ends of the wires; and b) the secondary travelling waves arise as a result of reflection of the above wave from the wire ends and move from there back toward the feed. Interference between the above waves, at the feed, produces the standard dipole's, undesirably large, variation-as-a-function-of-frequency characteristic. Travelling wave antennas represent an attempt to solve the above problem by the deliberate introduction of attenuation for the travelling waves and hence (at the feed): a) reducing the amplitude of the secondary travelling wave, leaving essentially b) the wave travelling away from the feed. It is possible to achieve this through the use of long antenna wire (e.g. the helical beam antenna) and relying on radiation to supply the required attenuation. This type of antenna is however too large for our purposes. In the antennas considered below, the above attenuation is obtained by: a) the introduction of resistors into the antennas structure; or b) placing the antenna wires sufficiently close to a non-perfectly conducting ground. Note that the antennas under a) and b) above tend to be inefficient, because power absorbed in the resistors or the ground represents power which is not radiated.

3.2.1 Travelling Wave Dipole

This antenna, is popularly known as the Australian dipole, since it was devised by engineers there [Altshuler, 1961; Guertler and Collyer, 1973; Treharne, 1983], and commercially developed by Antenna Engineering, Australia. Versions of it are however now available from several manufacturers (in the US and UK).

It is a multi-wire dipole antenna, to make it effectively "fatter" and so broaden the bandwidth, that is resistively loaded by inclusion of a resistor in series with each arm of the dipole (see Fig. 14a). The Australian version has an inductor across the resistor. The antennas we tested (the Australian and the UK versions) had series resistors of about 300-ohms, and a transmission line having an impedance of 300-ohms was used to feed the antenna (actually a 300:50 ohm balun was used). The dipole end provides a reactance against which the resistance is terminated, but this "termination" is only effective when the dipole end length is near to a quarter wavelength. Notwithstanding, the antenna exhibits a good SWR vs. frequency (nominally better than 2:1) over a broad bandwidth (say 3-30 MHz). No gain figures are given in the technical information available from the various companies that manufacture this antenna, but its radiation efficiency is claimed to be better than a terminated folded dipole (Fig. 14c). Royer [1987c] and Austin and Fourie [1988] have modelled the antenna, and given impedance, SWR, and gain. Royer has made the most detailed analysis, and his work considers the antenna pattern as well. There appear to be problems associated with achieving optimum NVIS gain over even the restricted bandwidth of 2 MHz to 10 MHz. Note that some companies claim a bandwidth of 1.5 MHz to 30 MHz for this type of antenna. There are problems with the azimuthal radiation pattern as well, which at the higher frequencies is more like a long wire pattern than a dipole pattern. Both of these researchers have independently suggested using two travelling wave dipoles on the same mast to improve performance.

Royer has written several documents [1985, 1987a, 1987b and 1987c] in which attempts were made to find an antenna which satisfies the following specifications:

a) The antenna is to serve as a part of short to medium distance communications system, employing frequencies in the 2 MHz to 10 MHz band, where the signal's propagation path includes one bounce off \bar{J} .

the ionosphere. It is desirable therefore that the antenna should have high gain in the vertical direction; and

b) The antenna's construction is to be such that it can be quickly moved from site to site. It was decided therefore that the antenna mast height should be no more than 15m.

Royer's 1987c document contains the results of an investigation to find a travelling wave drooping dipole antenna system which was suitable for the above application. The system shown in Fig. 15a was recommended. To obtain satisfactory performance across the band 2 MHz to 10 MHz, it was necessary to employ two travelling wave drooping dipoles with leg lengths of 60m and 30m operated over the bands respectively of 2MHz to 4.2 MHz and 4.2 MHz to 10 MHz. If the antennas are mounted on the same mast then (as shown in Fig. 15a) one of the antennas should be rotated 90° with respect to the other. It was found that, without this rotation, the presence of the 60m antenna unsatisfactorily changed the 30m antenna's gain characteristic. Both one-wire and two-wire dipoles were considered. It was shown that the input impedance (Z_a) values for both was high with respect to characteristic impedance of conventional transmission lines. The typical Z_a values for two-wire antennas were found to be significantly less than those for one-wire antennas. Therefore, for the purpose of easing the problems associated with matching the antennas to their transmission lines, the recommended system was constructed using two-wire dipoles. The variation of Z_a , across the frequency bands for the 60m and 30m dipoles was found to be least when the load resistors (R_L) had values of respectively 800 Ω and 1000 Ω . With these values for R_L , the VSWR on the feeding transmission lines (providing the antenna sees the lines characteristic resistances as about 700 Ω and 600 Ω for respectively the 60m and 30m dipoles) was found to be less than 2:1 across the recommended bands. In other words, it would not be necessary to tune these antennas. Fig. 15b shows vertical gain versus frequency for: a) the recommended travelling wave drooping dipoles, as computed using NEC-3 where $\epsilon_r=15$ and $\sigma=8\text{mS/m}$; and b) resonant half wave dipoles, 30 feet (9.1m) above the ground, measured over sandy soil where it was estimated that $\sigma=1\text{ mS/m}$. It would be desirable to devise a broad-band dipole that had the gain of a resonant dipole at the same height. The measured gains for the prototype TWD antennas are also graphed in the figure. Note that there is fairly good agreement between theory and experiment, excepting for a systematic difference with increase in frequency. This difference can be accounted for if the conductivity at the test site were lower than estimated. The ground conductivity should have been measured rather than estimated if closer agreement was sought. A part of the difference between the gain of the travelling wave dipole antenna and the resonant dipole is due to losses in the load resistances. The increasing difference at higher frequencies, near the upper end of the band for each of the two TWD antennas is due to difficulties associated with their patterns. At the low frequency end the gain is maximum overhead. As the frequency is increased the gain maximum shifts away from the vertical

3.2.2 Terminated Dipole

Another form of broadband dipole antenna employs direct resistive terminations to earth at the ends of a fan type drooping dipole see Fig. 14d [ref. Paterson, 1982]. Calculated patterns (directive gain in dBi for a perfect earth) reveal a rather omnidirectional pattern over the frequency range 2-10 MHz, with some lobing in the frequency range 10-30 MHz; for an antenna size, height 12m (40 feet), length 67m (220 feet) and width 17m (56 feet). The pattern is like that for a resonant drooping dipole, viz. high angle horizontal polarization in the plane broadside to the antenna and high and low angle vertical polarization in the plane of the antenna. Directive gains of 5 dBi are given in the above referenced technical note, but for proper evaluation gain or antenna efficiency is needed. It is not physically possible to design an antenna whose ends are near to the ground and terminated without incurring a significant loss in radiation efficiency. Efficiencies ranging from -13 dB at 2 MHz to -1 dB at 30 MHz are claimed for this antenna (presumably for perfect earth). Actual gains have been measured for comparison with other antenna types, see below.

3.2.3 Terminated Sloping Wire

Lt. Col. Alan Christinsin, USAF [1976] is the originator of the terminated sloping wire antenna, which he called the AFØNXX long wire. It was devised as a "quick and dirty antenna" to put up for tactical communications. Like the sloping V to be described, it is a sloping wire (see Fig. 16a) about 15.2m (50 feet) high at one end, 152m (500 feet long), and terminated to the earth at the far end (by about 600 ohms).

The long wire is end fed and hence must be operated in conjunction with a suitable earth ground. Since a part of the antenna runs parallel to the support mast, the mast should be non-conducting. The

input impedance of the wire is about 600 ohms, hence a 50:600 transformer is needed for 50 ohm coaxial feed.

The gain of the antenna is said to be better than the sloping V (to be described) on frequencies less than 9 MHz, and above 18 MHz it is worse. A disadvantage is that the direction of fire does not coincide with the direction of the wire, but is at an angle to the wire. This angle becomes smaller as the frequency increases. The direction of the wire needs to be off the great circle bearing by about 20° for the frequency band 4-6 MHz, 7° for the band 6-12 MHz and on bore site for the 12-18 MHz band. Also it is not a near vertical incidence radiator, so it will not be considered further here.

3.2.4 Sloping V/Sloping Triangle

The sloping V antenna (see Fig. 16b) is a resistance terminated antenna that was developed for tactical communications. In a typical configuration it is 15.2m (50 feet) high at one end, and each wire is 152.4m (500 feet) long. There are two wires and each is terminated in a 300 ohm resistor. In some versions the two ends are connected, and the termination is from the center of the base of this triangle (Fig. 16c). Apex angles of 32 - 50° have been employed. Both wires are fed, employing 600 ohm open wire line; or a 50:600 ohm balun could be used, mounted at the top of the mast, so that the antenna could be fed by coaxial cable.

Antenna gains specified by manufacturers of this antenna tend to be optimistic, e.g. a major antenna manufacturer (U.K. company) quotes gains of 13, 16, 18 and 21 dBi for frequencies 8, 12, 20 and 30 MHz. We shall see below that these gains do not take into account antenna efficiency. Theoretical gains have been calculated by Thomas and DuCharme [1974], Faust [1986] and by the Harris Corporation [private communication 1986]. The Thomas and DuCharme calculations were based on a simple theory. The calculations by Faust and by the Harris Corporation employed a more sophisticated theory (NEC-2). These gain values are plotted on Fig. 17. The calculations were made for poor and good ground, and there are other differences. The leg length of the sloping V for the Thomas and DuCharme model was 122m (400 feet); the leg length for the other models was 152m (500 feet). Notwithstanding, there seems to be fair agreement between the curves for frequencies below about 15 MHz, but for frequencies >15 MHz the Thomas and DuCharme gain values start to decrease. The predicted gain is a maximum when the leg length is equal to about 8-wavelengths. These calculations were for poor ground, whereas the other two curves, and the measured data point 4 [Faust, 1986] refer to a sloping-V over ground of good conductivity.

The curves of Fig. 17 are the gain in the direction of the pattern maximum, which is the bisector between the sloping arms of the antenna. The polarization is horizontal. The vertical take-off angle varies from about 30° at 4 MHz to 10° at 30 MHz. There are however two major lobes on each side of the direction of fire, in the direction of the arms of the antenna (the long wire pattern). The radiation in these directions is vertically polarized. The curves on Fig. 18 are measured data, at an elevation angle of 10° for an antenna over good ground at 22.9 MHz.

3.2.5 Half Rhombic(Inverted-V)

The half-rhombic, sometimes called an inverted V travelling wave antenna is essentially one-half of a rhombic antenna split along its major axis and turned so that its plane is vertical, see Fig. 19a. In this configuration the main lobe is vertically polarized. The construction is simple, since it requires only one supporting pole. The ground forms one side of its input and output circuits. The antenna is terminated at one end (by about 300 ohms) and fed at the other end.

Like the rhombic the antenna needs to be long, since its radiation efficiency drops off rapidly when the leg length (the length of the sloping wire on each side of the mast) becomes less than a wavelength. Predicted gains [Thomas and DuCharme, 1974] for this antenna over poor ground, employing a radial ground screen at both ends, are -1 and 3 dBi when the leg length is 0.5 and 1λ . The gain rises quickly with increase in frequency (leg length becomes long wrt a wavelength), and when the leg length is 2.5λ the predicted gain is 10-12 dBi. It is essentially a low angle radiator, when used at frequencies for which it is an efficient radiator, but compact versions of it have been used for NVIS links. Royer [1987d] has calculated the gain and pattern for a compact size inverted-V, and the performance of a commercial version of such an antenna has been measured (see below).

3.2.6 Delta

The delta is a wire antenna, supported by a single mast, having triangular shape, see Fig. 19b. The apex of the triangle is up, and it is fed in the middle of the base with 600 ohm line. The antenna is terminated at its apex in 600 ohms.

It is a useful wide-band antenna for short to medium distance communication, but its efficiency for a 15m (50 feet) mast height is very poor. It is therefore not a very suitable antenna for the tactical communications requirement. Large size versions of it have been used for years for ionospheric sounding, however, excepting for the simplified analysis of Bailey [1951], little attention has been given to evaluating its performance. A tandem delta has been developed by Antenna Engineering, Australia (Fig. 19c). This antenna comprises two orthogonal deltas, one terminated in the other. It provides a claimed 3 dB more gain, since there is no terminating resistor. The pattern, like the delta is essentially high angle, but the azimuthal pattern is said to be more omnidirectional compared with the standard delta. Both versions have been numerically modeled by Royer [1985], and performance has been evaluated by measurement (see below)..

3.2.7 Antennas closely-coupled to the Ground (Beverage Type)

The **Beverage antenna** for use at HF is well described in two recent articles, one co-authored by the originator of the antenna Beverage and DeMaw [1982]; and one by author Belrose et. al... [1983].

The Beverage is a good low noise receive antenna, which, if one has sufficient real estate, can provide good directivity at low HF frequencies. It is used mainly for point-to-point communications, however, a steerable version has been described [Misek, 1977]. The efficiency of a Beverage antenna depends on its length, and on ground conductivity. At frequencies near the low end of the band of interest (2MHz) its gain decreases as earth conductivity increases, whereas at higher frequencies this variation has the opposite trend. Gain increases with increase in length of Beverage antenna, until its length exceeds about 4λ , for a Beverage above average ground..

The pattern is suited for long distance sky-wave communications, since the launch angle is low, and for short/medium distance ground-wave links. If used as a transmitting antenna several Beverages must be fed in phase, since the efficiency of a single Beverage element is poor. Because the efficiency is low the individual elements can be quite closely spaced, since coupling is small, and the gain is approximately doubled each time the number of elements is doubled.

The Beverage antenna belongs to a class of antennas that are closely coupled to the ground, in fact performance relies on the ground to enhance the received signal. The antenna concepts employed in the design can be viewed as synergistically interacting with the ground environment. Another antenna type belonging to this family has recently been devised by the Eyring Research Institute, dubbed the **Eyring Low Profile (HF) Array (ELPA)** [Faust and Skousen, 1986]. This antenna is attracting a considerable military interest since the commercial version is lightweight, easily deployable, and requires no support mast or tower (see Fig. 20). It is a wide band antenna system that provides bi-directional end-fire directivity (although it can be configured as a Beverage), suited for long distance communications for frequencies for which it is electrically long (say >10 MHz). At frequencies near the low end of the band its pattern reverts to a high angle radiator for sky-wave, but it also radiates ground-wave in the end-fire directions.

The efficiency of a single element, like the Beverage is rather poor. However, as is the case for inefficient antennas, the array gain is approximately doubled (3 dB) each time the number of elements is doubled. Thus the 4-element array in Fig. 20 has about 5 dB more gain than does a single element.

At the time of writing, one of us (author Royer) was in the process of computing (using NEC-3) the characteristics of single wire versions (as shown in Fig. 21a) of the ELPA type of antenna. The vertical gain for this array was found to have two sharp dips at 8.4 MHz and 21 MHz (Fig. 21b). As the vertical patterns in Fig. 21c illustrate, these dips in gain are not caused by increased losses in the ground, but instead arise because the gain pattern maximums have shifted away from the vertical.

4. Measured Performance of some commercial HF antennas

The relative gain of an antenna under test was measured over a near vertical incidence sky-wave (NVIS) link (100 km distant) with reference to a resonant horizontal dipole at 9.1m (30 feet). For transmitting the HF signals, horizontal half-wave dipole antennas were employed, arranged in line to minimize the number of support poles and minimize coupling between antennas. All antennas were arranged in such a direction to have them aligned broadside to the receiving antennas. A similar arrangement for the reference dipoles was employed at the receiver location.

Measurement of the strength of the RF signal received on a calibrated FI meter from the reference dipole and the antenna under test were conducted over a three to five minute period, when the short term fading rate was not greater than a fade about every 20-seconds. A coaxial switch enabled the selection of the signal from either the reference or antenna under test. In order to minimize errors when the signal is fading rapidly the signal strength expressed in dB above 1-microvolt was alternately measured on the reference antenna and the antenna under test by recording the signal just prior to switching and just after switching. The difference in signal strength between the antenna under test and the reference dipole antenna was computed from at least 30 measurements, and a median value determined. The median value was rejected and the measurements repeated if the upper and lower quartile values differed by more than 1 dB from the median value. Account was taken of the different lengths and attenuations of the coaxial cables used with the various antennas.

The ground conductivity at the receive site was estimated, based on experience, and using this conductivity value the gain of the reference dipole antenna at the three frequencies used for the measurements was calculated (using NEC). Using this calculated gain, the actual gain (in dBi) of the antenna under test could be estimated.

This gain vs frequency for a number of commercially available antennas (ranging in price from a few hundred dollars to a few tens of thousands of dollars) and several CRC prototype antennas are co-located on three graphs, Figs. 22, 23 and 24. The first graph collects together the results for tuned/or resonant antennas. The second graph collects together the results for resistance loaded travelling-wave type of antennas. The third graph collects together the results for delta antennas and for several miscellaneous types of antennas

As anticipated at the outset, resonant dipoles exhibited the highest gains. The gain of a drooping dipole was insignificantly different from that of a horizontal dipole. The performance of the entire class of travelling wave antennas was relatively poor, some were very much inferior to others. The best of the resistive/or resistive and reactive loaded antennas, the travelling wave dipole and the large delta antenna on a 30m (100 foot) mast were 3-4 dB down on the resonant drooping dipole antenna. The efficiencies of some of the antennas tested fell with the square or the fourth power of the working frequency, and some were more or less radiating dummy loads at frequencies near the low end of the band (say 2 MHz). All of the travelling wave type were too small to yield an acceptable gain at 2MHz. As noted by Royer [1987c] a travelling wave dipole about twice as long (60m) is required, and such a dipole is too long to provide optimum NVIS gain for frequencies > 4MHz. In fact Royer suggested the use of two travelling wave dipoles, one 60m long, and one 30m long to optimize performance over the whole of the band of interest (2-8 or 2-10 MHz)..

5. On Ground Systems

Since ground forms one side of a monopole antenna, some form of ground system must be a part of the antenna system. For mobile and transportable application this ground system could be the vehicular ground, a ground stake, or a radial wire system laid on the surface of the ground or elevated at a small height above the ground. For a radial ground system laid on the ground insulated wire or bare wire could be employed.

Breakall [1986] has extensively investigated (by numerical modelling using NEC-2) the effect on impedance and gain for a quarter wave monopole over various types of ground, using a ground stake (2-, 4- and 6-foot lengths); or radials (0.25, 0.4 and 1λ long), buried or elevated. The number of radials was also a parameter (4- to 360-radials).

In Figs. 25 and 26 we have plotted some of the numerical results of Breakall's study. In Fig. 25 we have plotted the theoretical gain (dBi) for a monopole antenna vs. number and length of buried radials

for three ground conductivities (very poor, average and very good ground). The heavy lines and closed symbols correspond to $\lambda/4$ -radials; the lighter lines and open symbols correspond to radial lengths λ . Triangles correspond to very poor ground ($\sigma=1.5 \times 10^{-4}$ S/m); squares to average ground ($\sigma=3 \times 10^{-3}$ S/m); and circles to very good ground ($\sigma=5 \times 10^{-2}$ S/m). The curves in Fig. 26 compare elevated radials with buried radials. In this figure we have plotted gain vs. ground conductivity. The three lower curves are for elevated $\lambda/4$ radials for three heights ($h/\lambda=0.0033$, 0.033 and 0.1 respectively). The top curve is for 120 $\lambda/4$ radials buried. A close inspection of these graphs reveals that:

- 1) For buried $\lambda/4$ radials, the increase in gain for $n>30$ is marginal. The maximum gain is a strong function of ground conductivity;
- 2) A large number of long radials (λ in length) can compensate to some extent for poor ground conductivity. The launch angle ψ is however a function of ground conductivity ($\psi=20^\circ$ for good ground, $\psi=30^\circ$ for poor to very poor ground). The gain values plotted on Figs. 25 and 26 are the maximum values; and
- 3) When the ground conductivity is very poor a large number of radials should be used; excepting if there are few radials ($n<30$), quarter wavelength radials should be used.

Christman [1988] has published some of the results of this study, and he has extended the work, particularly concerning the use of elevated radials. Elevated radials, from an electrical point of view are preferred to buried radials, or to radials lying on the ground. The efficiency of a monopole with four elevated radials (Fig. 26), even one-radial (see below) is not significantly different from a monopole with 120-radials buried. For the Canadian environment, however, the cost of installing elevated radials, for fixed location operation, is expensive due to ice loading.

An example an interesting antenna is discussed below. Since the military communicator in a tactical situation may want to communicate to both short and long distances, it is sometimes useful to employ an antenna that radiates a mix of vertical and horizontal polarization. A simple wire antenna that does this has been devised by author Belrose. The antenna is a dipole type, wound up on bobbins so that it can be unrolled to the right length for the frequency used (each arm is a quarter of a wavelength long). For transportable application the support mast should be a telescopic insulated type, of height at least 15m. For short to medium distance communication links the antenna would be configured as a drooping dipole (Fig. 27a). For communications to greater distances, and for ground wave links the antenna would be configured as a monopole with one elevated radial (that is the two arms are at right angles to each other). This configuration (Fig. 27b) which has been referred to in recent radio amateur literature as the **VE2CV Field-Day Special** was also modelled by Christman [1986]. The radial would point in the direction that one wishes to communicate, but only approximately since the azimuthal pattern is wide. The antenna exhibits a front-to-back ratio of 12-15 dB at some take-off angles (see Fig. 28); the input impedance is 50 ohms; and it produces both low and high angle radiation. The maximum gain is on par with a 120-buried-radial vertical monopole. As is the case for any vertically polarized antenna the signal on the horizon is not a null, the figure shows the sky-wave pattern. The signal on the horizon is ground-wave.

The pattern for the reference antenna has been calculated for average ground ($\sigma=3$ mS/m). As the electrical quality of the soil becomes worse the height of an elevated radial system must be raised to progressively higher heights above the earth in order to realize performance on par with a 120-buried radial vertical monopole. At 1 MHz Christman's studies have shown that adequate heights are 3m for good soil, 5m for average soil and 7m for very poor soil. This corresponds to 0.01 , 0.016 and 0.023λ . Therefore if high angle radiation is not required, the VE2CV Field-Day Special antenna could be operated quite close to the ground without loss in gain (height of radial 1-2m for a frequency of 3.8 MHz over average soil).

6. Concluding Remarks

This lecture is based on the considerable combined experience of the authors with modelling, fabricating, measuring performance of and using practical HF antennas for the low HF band (2-8 MHz). This is an important frequency band, which is used to provide short to medium distance point-to-point, mobile/transportable radio communications for military and civilian applications. The antenna poses a particularly severe problem, particularly if one wants a broad-band antenna for frequency agile systems, because of the four-to-one change in wavelength over the band.

The best simple antenna for the requirement where only a few frequencies or channels are used is a drooping dipole array, where several drooping dipoles are connected in parallel with a common feeder. The drooping dipole arrangement makes for a very practical and efficient antenna system. The dipoles can be fanned (angular separation) in a horizontal or a vertical plane. However there is some interaction between the dipoles, and a very low VSWR (say $< 1.5:1$) is not always possible. The situation where one dipole is a third harmonic of a lower frequency dipole should be avoided. If a greater number of frequencies are operationally used (say four or more), or if complete frequency agility is required a dipole with an automatic antenna system tuning unit (ASTU) will provide the best radiation efficiency. However this complicates the antenna system, and (undoubtedly) compromises reliability. A switched drooping dipole antenna has been studied, but the "switching" (i.e. band change) for the prototype fabricated was done manually, by lowering one arm of the dipole at a time. A staggered tuned dipole (comprising several dipoles whose lengths are logarithmically spaced across the desired band), or a bow-tie dipole, both types employing an ASTU, can be designed for specific requirements.

Resistance loaded travelling wave antennas are extensively employed for frequency agile systems. Clearly such a dipole is a compromise, since there are problems with gain and pattern. We have measured the performance of many of these, if not all of the travelling wave antenna types so far devised. The best performer (for NVIS gain) is the delta antenna, which is why it has been so extensively used for vertical incidence ionospheric sounding; but it has to be physically large, employing a 30m high (or higher) support mast if acceptable performance at say 2 MHz is desired. The small delta on a 15m mast is a disaster. The best performer for low heights (9-15m) is the travelling wave dipole antenna. This antenna has been around for quite a while, but a great deal of effort has been made in recent years to understand how it works, and to improve its performance (including studies by the authors).

While we have not addressed the subject of baluns (balanced-to-unbalanced transformers) we have noted that there is a need to develop a **balanced** automatic microprocessor controlled antenna system tuning unit. Conventional baluns work best and should only be used when the VSWR is not too high. The best balun in the author's view is a ferrite bead balun [Maxwell, 1983], at least for nominal transmitter powers (a kW or so), since no transformer winding is required.

We have been particularly concerned with antennas at low heights, even with antennas on the ground. Antennas lying on or near the ground are of particular interest for military tactical communicators, since such an arrangement is a low profile antenna system, and no support mast(s) is required. HF radio operators in the Canadian north have frequently reported success with using HF dipoles lying on the ground, hence undoubtedly such operational configuration work. They would work even better if their length were right for resonance, or if they were used with an ASTU. The reason that antennas perform at all under such circumstances is because of the very low ground conductivities found in the Canadian north. This limits the application of such low profile dipole systems, since they should not be used at locations where the ground conductivity is high. Their performance would be a disaster if laid on sea ice, and not very good if laid on fresh water lake ice. In this lecture we have shown (for the first time in the authors experience) what gain to expect under such circumstances.

We have particularly tried to compare computer simulation with measurement, however the comparisons made were not all as satisfactory as we would have liked, either because we did not know the ground conductivity at the test site (at least not exactly), or because calculations for the right conductivity were not available at the time of writing.

Acknowledgement

The authors would like to acknowledge the technical assistance of P.R. Bouliane and E.E. Stevens. Mr. Bouliane fabricated the various CRC prototype antenna systems, and made impedance measurements for these and for dipole antennas vs height. Mr. Stevens assisted author Petrie with the set up of the many antenna systems evaluated and tested, and with making the antenna gain measurements. The authors are indebted to the Department of National Defence, who funded a study, which provided the inspiration and basis for this lecture.

References

- Altshuler, E.E. [1961] The travelling-wave linear antenna, IEEE Trans. Antennas and Propagation, AP-9, 324-329
- Austin, B.A. and A.P.C. Fourie [1988] Numerical modelling and design of loaded broadband wire antennas, IEE Conf. Proc. 284, 4th International Conf. on HF Systems and Techniques, 125-129, London.
- Bailey, R. [1951] Aperiodic aeriels use with vertical incidence ionospheric recorders, Wireless Engr. 28, 208.
- Belrose, J.S., J. Litva, G.E. Moss and E.E. Stevens [1983] Beverage antenna for Amateur Communications, QST, January 1983, pp. 22-27 (correction QST, May 1984, p.46).
- Beverage, H.H., and D. DeMaw [1982] The Classic Beverage Antenna, Revisited, QST, January 1982, pp. 11-17.
- Bhattacharyya, B.K. [1963] Input resistances of horizontal electric and vertical magnetic dipoles over homogeneous ground, IEEE Trans. Ant.&Prop., AP-11, 261-266.
- Breakall, J. [1986] Private communications.
- Burke, G.J. [1980] The Numerical Electromagnetics Code (NEC), in Applications of Moment Methods to Electromagnetic Fields (edited by Bradley J. Straite), The SCEEE Press.
- Christinsin A.S. [1976] Technical Report Frequency Management Digest. Anthology 1970-1976, Spectrum Management Division US Air Force Communications Service.
- Faust, D.L. [1986] Selected pattern measurements for four full scale Tactical HF Antennas, 2nd Annual Review in Progress in Applied Computational Electromagnetics, Monterey, CA
- Faust, D.L. and O.N. Skowsen [1986] Gain and pattern measurements of a low profile HF antenna array, Applied Computational Electromagnetics Society Newsletter, 1(No. 2), 31-46.
- Fitzgerrell, R.C. [1967] The gain of a horizontal half-wave dipole over ground, IEEE Trans. on Ant. and Prop., pp 569-71.
- Friis, H.T., C.B. Feldman and W.M. Sharpless [1934] The determination of the direction of arrival of short radio waves, Proc. IRE, 22, 47-78.
- Guertler, R.J.F. and G.E. Collyer [1973] Improvements in travelling wave dipoles, Proc. IREE Convention, Melbourne, 70-71.
- Hagn, G.H. and G.E. Barker [1970] Research Engineering and Support for Tropical Communications, Final Report, Stanford Research Institute, Menlo Park, CA 94025.
- Maxwell, W. [1983] Some aspects of the Balun Problem, QST, 38-42, March, 1983.
- Miller, E.K., A.J. Poggio, G.J. Burke and E.S. Selden [1972] Analysis of wire antennas in the presence of a conducting half-space. Pt. II. The horizontal antenna in free space, Can. J. Phys., 50, 2614-2627.
- Misek, V.A. [1977] The Beverage Antenna Book, Hudson NH: V.A. Misek.
- Paterson, J.P. [1982] Electrically-small HF transmitting Antennas, TCI Tech. Note 11, Technology for Communications International, Falls Church, VI.
- Patterson [1967] Electronics p. 111, 21 August, 1967.
- Proctor, R.F. [1950] J. Instrum. Elect. Eng. (London), 97, 133.
- Royer, G.M. [1985] A comparison of the characteristics of several wire antennas for operation in the 2-10 MHz frequency band, Communications Research Centre Tech. Memo. DRC 85-04, Ottawa.

Royer G.M. [1987a] The characteristics of a broadband switched drooping dipole, Ibid, Tech. Memo. DRC 87-03.

Royer, G.M. [1987b] Characteristics of two broadband (2-10 MHz) switched vee antennas, Ibid, CRC Report No. 1416.

Royer, G.M. [1987c] The characteristics of a broadband (2-10 MHz) travelling-wave drooping-dipole antenna system, Ibid, Tech. Memo. DRC-87-05.

Royer, G.M. [1987d] The characteristics of a terminated inverted-V over a frequency band 2-10 MHz, Ibid, Tech, Memo. DRC 87-04.

Royer, G.M. [1989] The bandwidth of a Dipole antenna in Free Space, Ibid, Tech, Memo. DRC 89-1

Stark, A. [1982] HF Antenna System AK 501 for mobile use, News from Rohde & Schwarz, 96, pp 12-16.

Thomas, J.L. and E.D. DuCharme [1974] HF Antenna Book: Calculated radiation Patterns, Communications Research Centre Report No. 1255.

Treharne, R.F. [1983] Multipurpose wholeband HF antenna architecture, J. Electrical and Electronic Engineering (Australia), 3, 141-152.

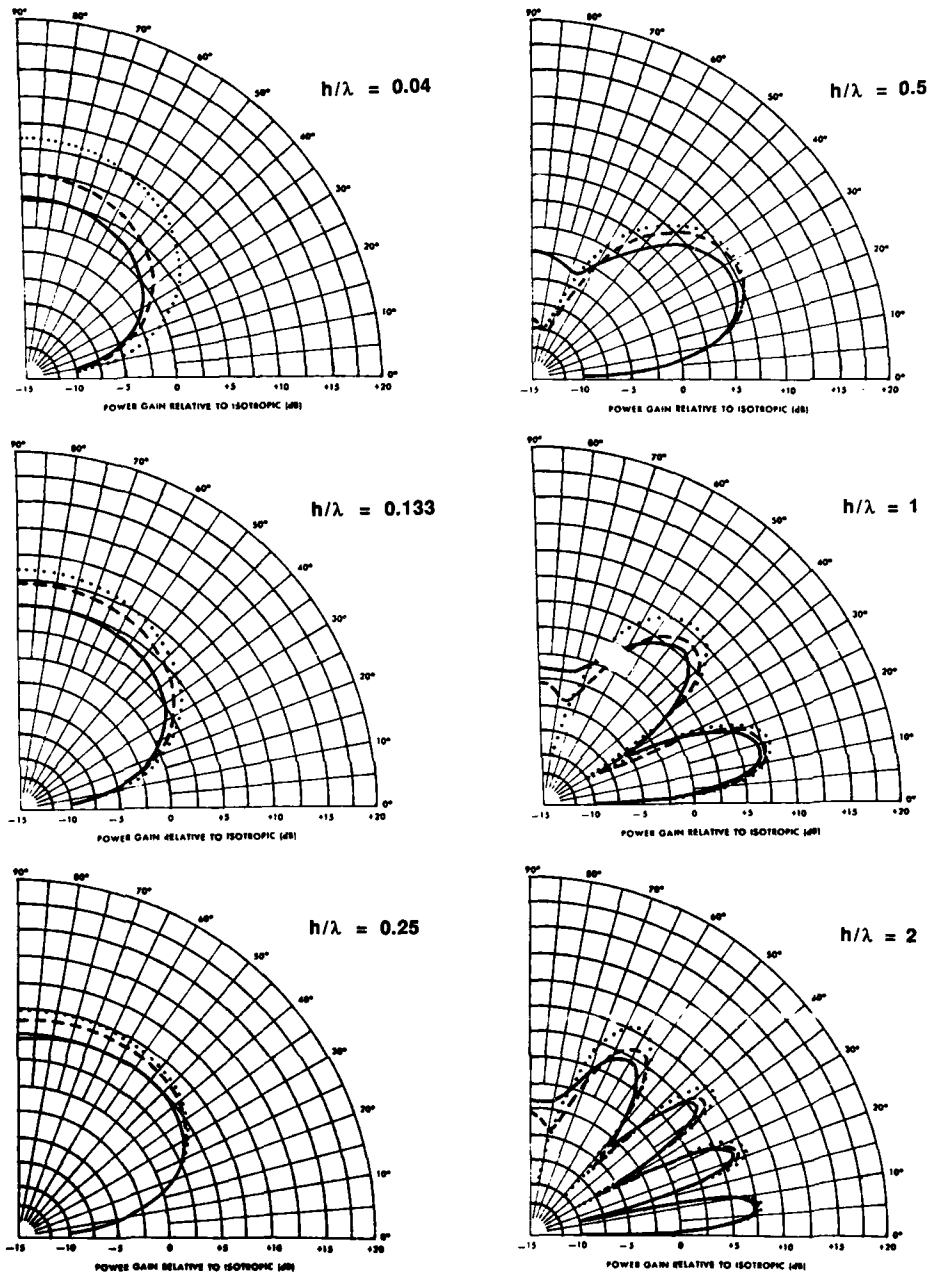


Fig. 1 Radiation pattern in the equatorial plane (early theory) for a horizontal half-wave dipole over three types of ground (medium dry ground: solid line; wet ground: dashed line; and sea water: dotted line) vs height of the dipole (h/λ) [after DuCharme and Thomas, 1974].

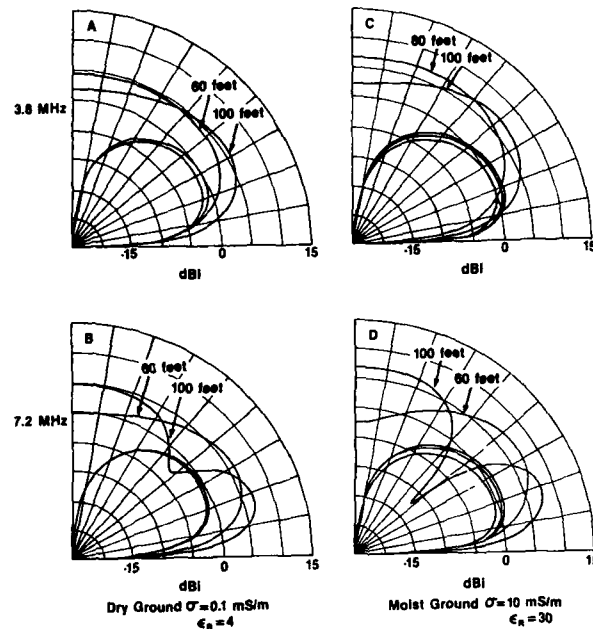


Fig. 2 Radiation pattern in the equatorial plane for a horizontal dipole at two heights (60 and 100 feet), two ground conductivities (dry and moist ground) and two frequencies (3.8 and 7.2 MHz); compared with the elevation plane response for a quarter-wave vertical monopole with radials a quarter wavelength long [after Breakall, 1986].

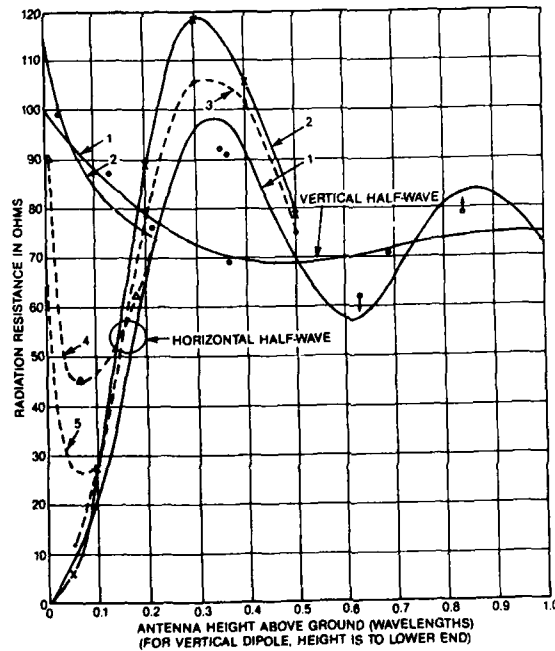


Fig. 3 Experimental and theoretical (early theory) impedance (resistive component) of a horizontal and a vertical half wave antenna over perfect and finitely conducting ground. Distance to dipole above ground in wavelengths (for vertical dipole distance is to the lower end).

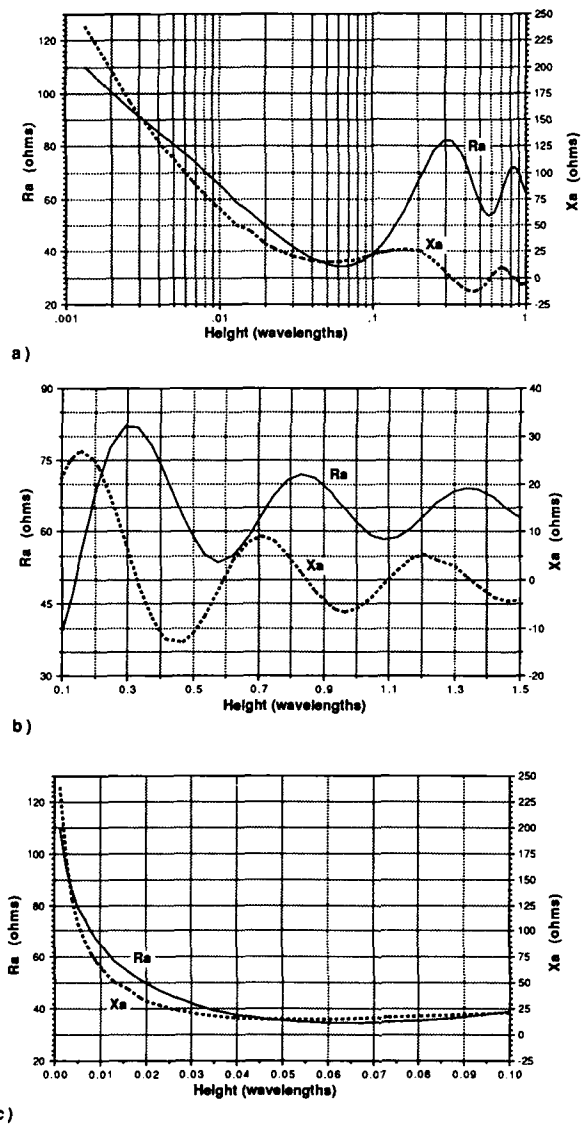


Fig. 4 Input impedance vs height (h/λ), according to NEC-3, of a 4 MHz horizontal dipole antenna (constructed of No. 12 antenna wire). Ground conductivity parameters: $\sigma = 10$ mS/m, $\epsilon_r = 1.5$.

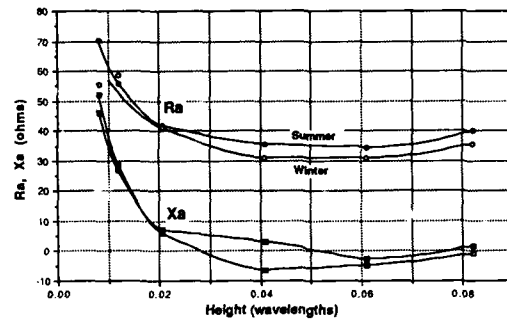


Fig. 5 Input impedance of a 4 MHz horizontal dipole vs height of the dipole (h/λ). The solid symbols are the measured values in summer; the open symbols the winter values (ground frozen and covered with snow). The ground conductivity at the test site, based on relatively crude ground wave propagation measurements, was estimated to be 10 mS/m (summer value).

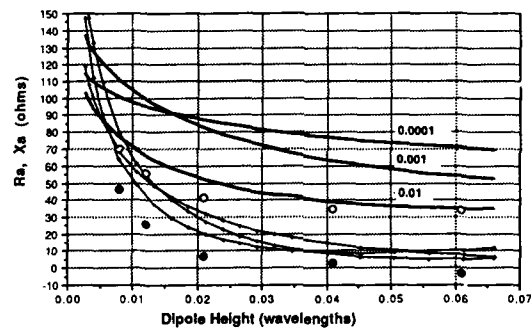


Fig. 6 Comparison between measured dipole impedances (open and closed circles), summer values and theoretical impedances, according to NEC-3; vs dipole height h/λ . The ground conductivities as marked are S/m (see table).

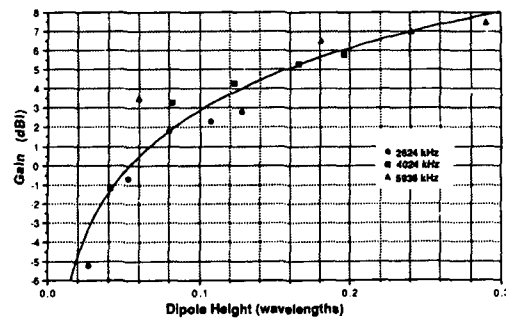


Fig. 7 Experimental measurements of near vertical incidence skywave (NVIS) gain (dBi) for horizontal half wave dipole antennas vs height of the dipole in wavelengths (h/λ); at three frequencies: 2624 kHz (circles); 4024 kHz (squares) and 5936 kHz (triangles). Ground conductivity at test site 1 mS/m (estimated).

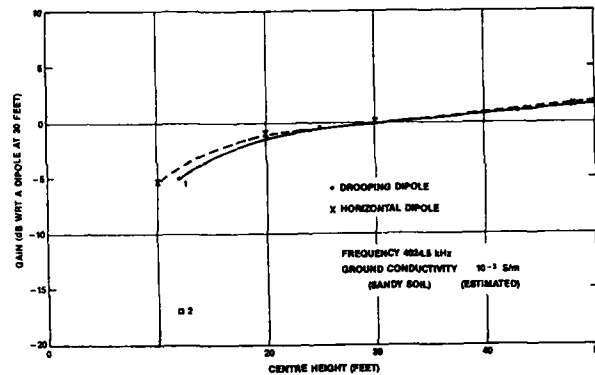


Fig. 8 Comparison between measured near vertical incidence skywave gain (relative gain wrt a dipole at 30 feet or 9.1m) vs height (feet) for a horizontal and a drooping dipole over poor ground (sandy soil, conductivity σ estimated to be 1 mS/m).

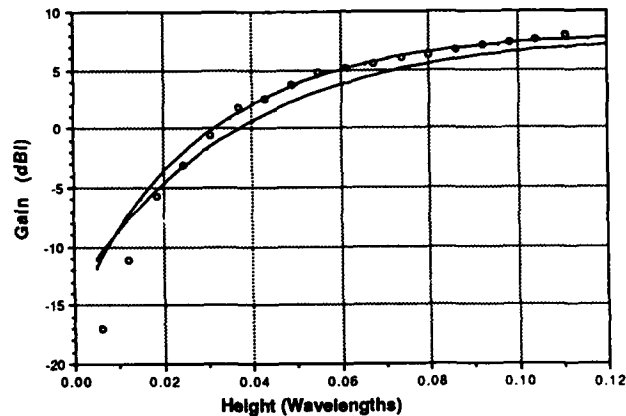


Fig. 9 Comparison between measured (data points plotted as an open circle) and calculated gains for a 6 MHz horizontal dipole vs height of the dipole (h/λ). The curves are the theoretical gains (according to NEC-3) for two ground conductivities, measured by the authors at the test site where they made their gain measurements, on two different occasions [Hagn and Barker, 1970]. The conductivity corresponding to the lower curve was 24 mS/m ($\epsilon_r = 51$); for the upper curve 58 mS/m ($\epsilon_r = 53$).

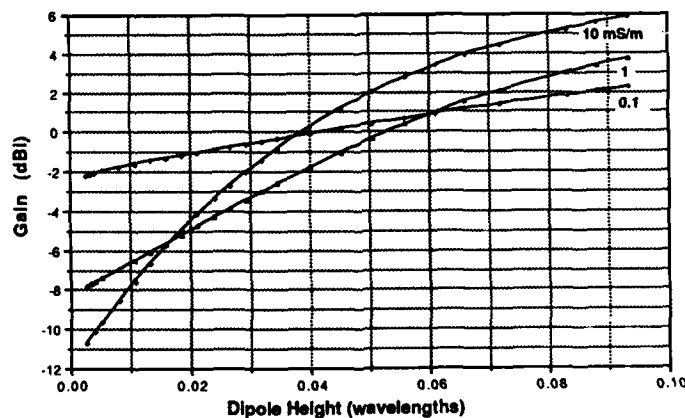


FIG. 10 Theoretical gains, according to NEC-3, for a horizontal dipole vs height of the dipole (h/λ) for three conductivities as marked (mS/m).

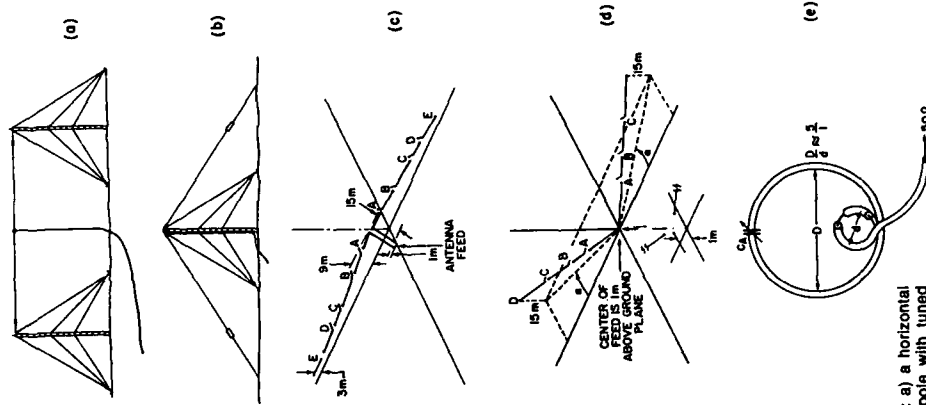


Fig. 12 Sketches of various types of resonant antennas: a) a horizontal dipole; b) a drooping dipole; c) a switched drooping dipole with tuned feeders; d) a switched Vee dipole; and e) a mini- or compact-loop.

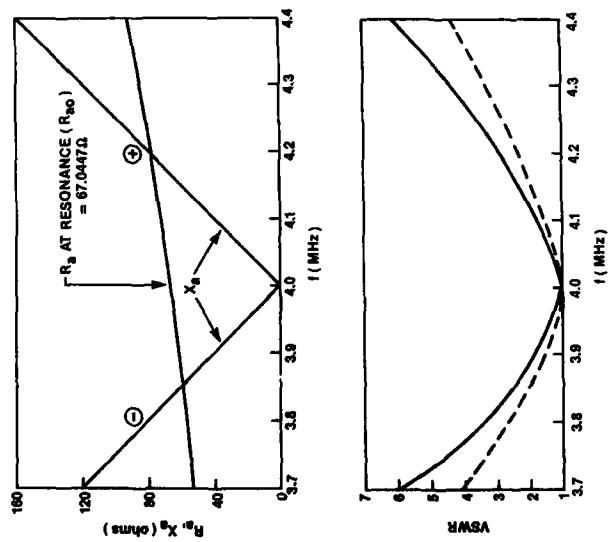


Fig. 11 Theoretical impedance, according to NEC-3, and VSWR for a resonant half wave dipole in free space, for a narrow band of frequencies near to its resonant frequency. The dashed curve (lower figure) is the measured VSWR for a drooping dipole at 50 feet (15.2m).

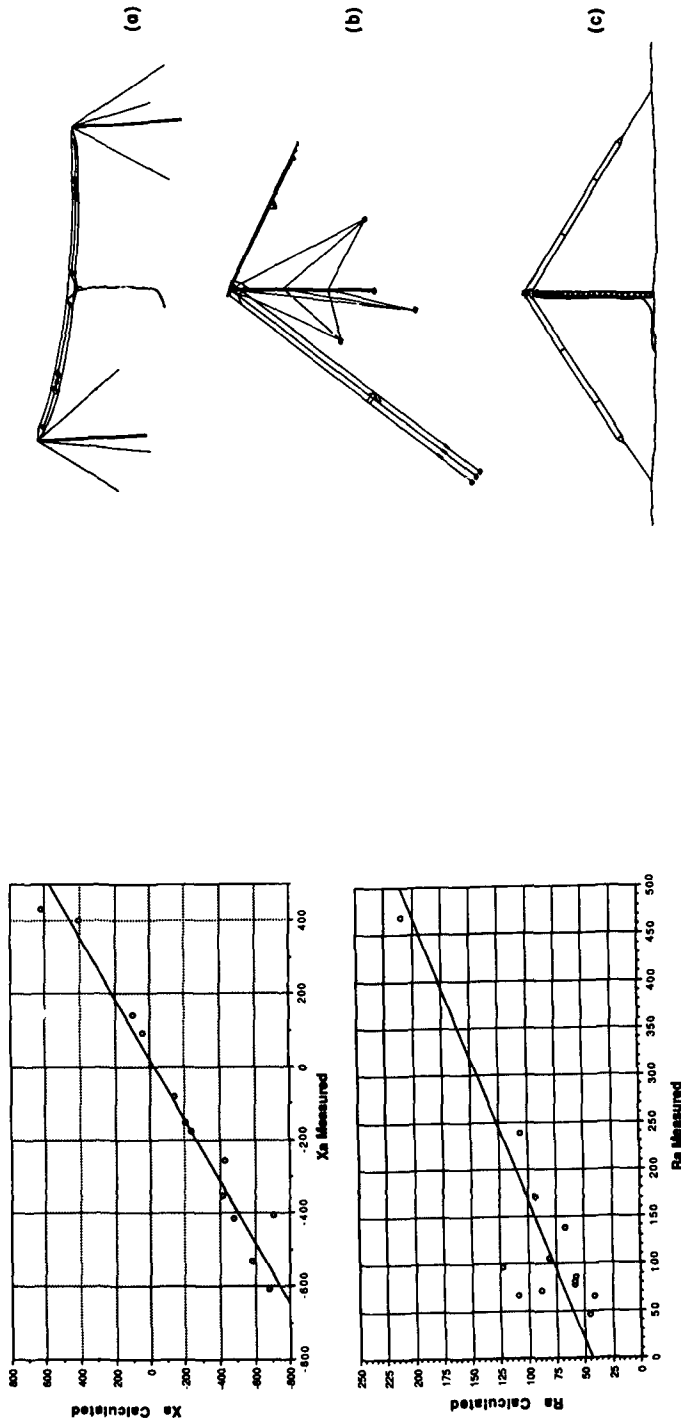


Fig. 13 Correlation between measured and calculated input impedance, at the input terminals of the 600 Ω transmission line feeding the CRC prototype switched drooping dipole antenna. A one-to-one agreement is not expected because the ground conductivity parameters taken for the calculated values was $\sigma = 8$ mS/m; $\epsilon_r = 15$; the measured values were for an antenna over sandy soil ($\sigma \approx 1$ mS/m).

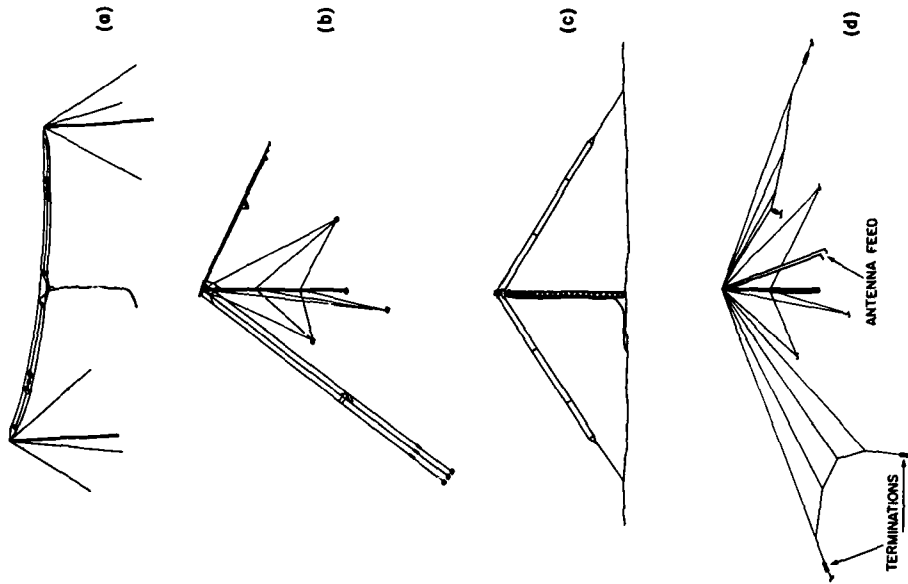


Fig. 14 Sketches of various configurations of travelling wave dipole type antennas: a) horizontal travelling wave dipole antenna (Australian type); b) a drooping travelling wave dipole antenna; c) a drooping terminated folded dipole; and d) a grounded dipole (ends terminated to ground).

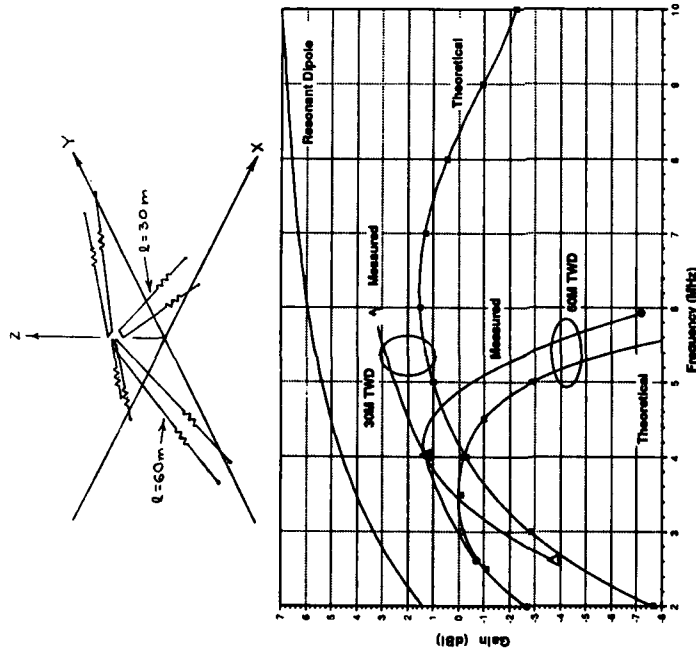


Fig. 15 Comparison between theoretical and measured gains vs frequency for the CRC prototype TWD antennas. The calculated vertical incidence gains according to NEC-3 are for ground conductivity parameters: $\sigma = 8\text{ mS/m}$, $\epsilon_r = 15$; and for load resistors R_L of 800Ω and 1000Ω for the $l = 60\text{ m}$ and $l = 30\text{ m}$ TWD antennas respectively. The measured NVIS gains are for TWD antennas with load resistors $R_L = 1000\Omega$; ground conductivity at the test site estimated to be 1 mS/m .

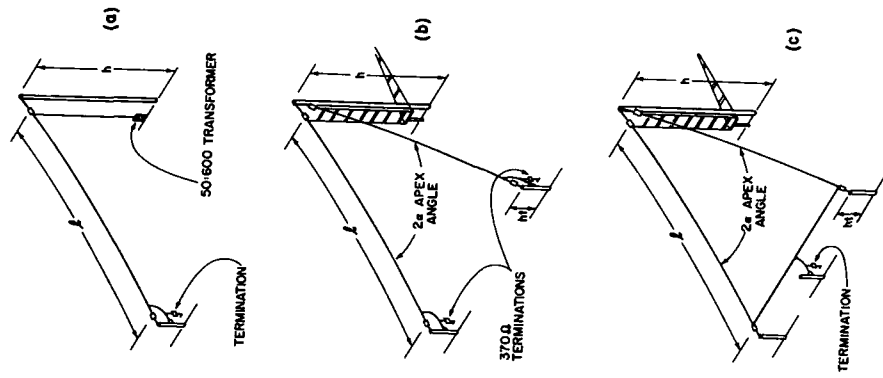


Fig. 16 Sketches of various types of long wire type of travelling wave antennas: a) a terminated sloping wire (AFØNXX long wire); b) a sloping v; and c) a sloping triangle.

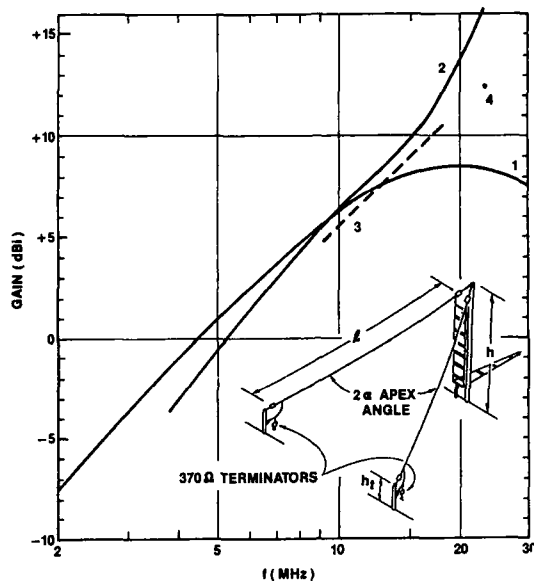


Fig. 17 Theoretical and measured (one data point) gain of a sloping-V over real earth (gain at pattern maximum).

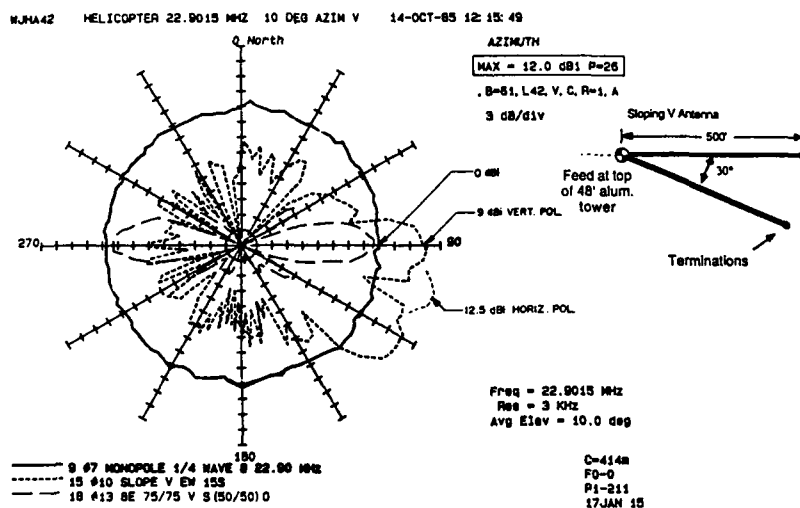


Fig. 18 Measured gain for a 500 foot sloping-V over good ground, for both horizontal and vertical polarization. Elevation angle $\psi=10^\circ$, frequency 22.9 MHz [after Faust, 1986].

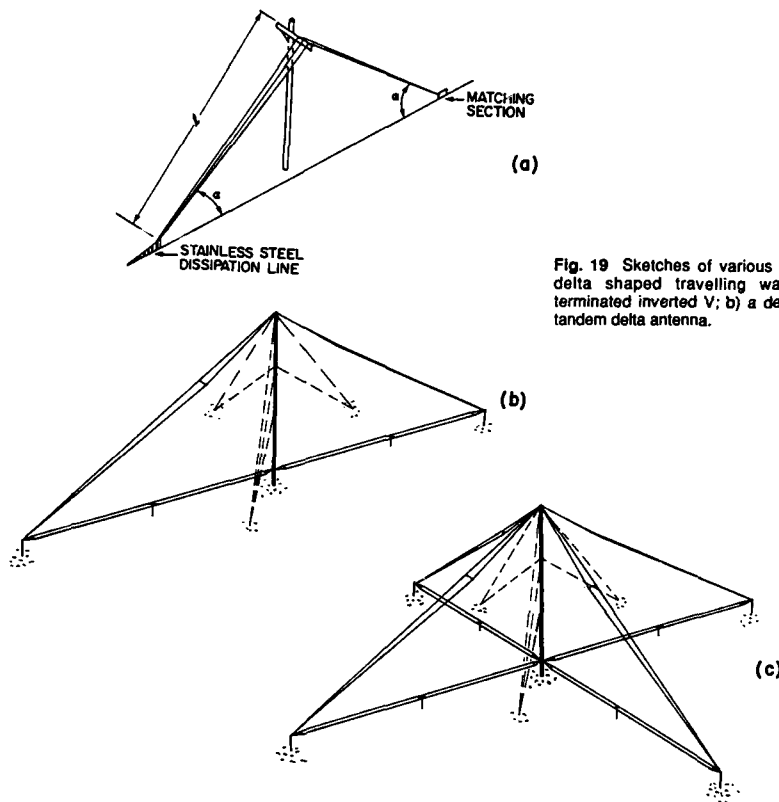


Fig. 19 Sketches of various types of triangular or delta shaped travelling wave antennas: a) a terminated inverted V; b) a delta antenna; and c) a tandem delta antenna.

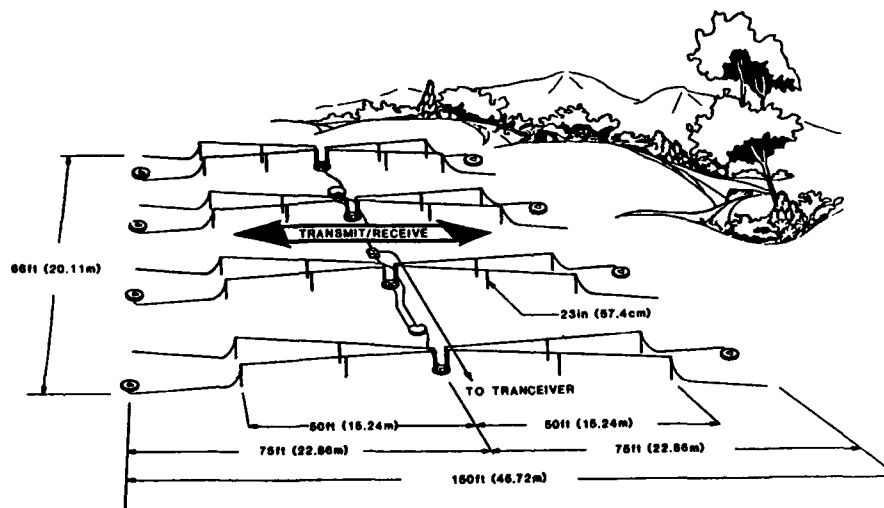


Fig. 20 Sketch illustrating the Eyring Low Profile HF Antenna Array (ELPA) [after Faust and Skowsen, 1986].

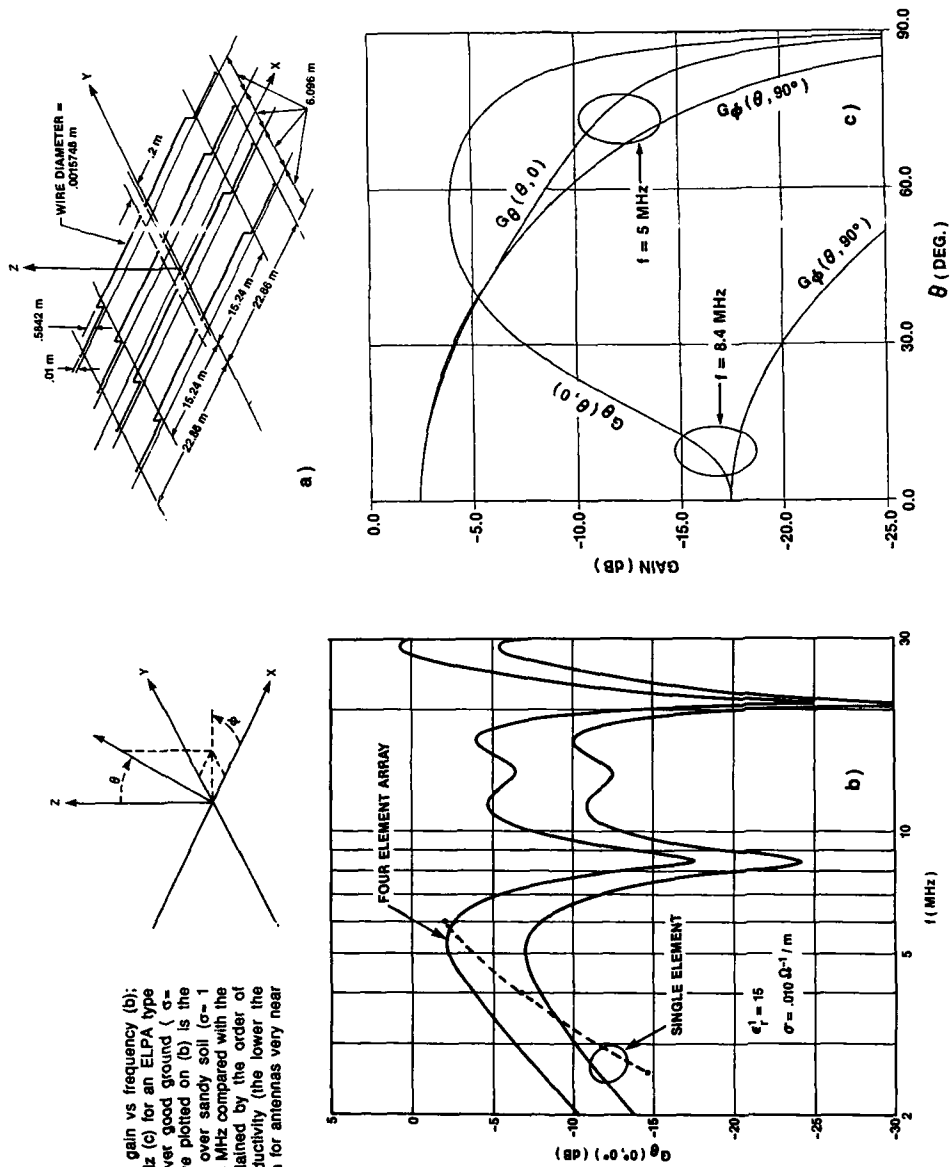


Fig. 21 Theoretical vertical incidence gain vs frequency (b); and radiation pattern at 5 and 8.4 MHz (c) for an ELPA type antenna array as sketched in (a), over good ground ($\sigma = 10mS/m$; $\epsilon_1' = 15$). The dashed curve plotted on (b) is the measured gain for a single element over sandy soil ($\sigma = 1 mS/m$). The higher measured gain at 6 MHz compared with the theoretical gain can in part be explained by the lower the magnitude difference in ground conductivity (the lower the ground conductivity the higher the gain for antennas very near to the ground, see Fig. 10).

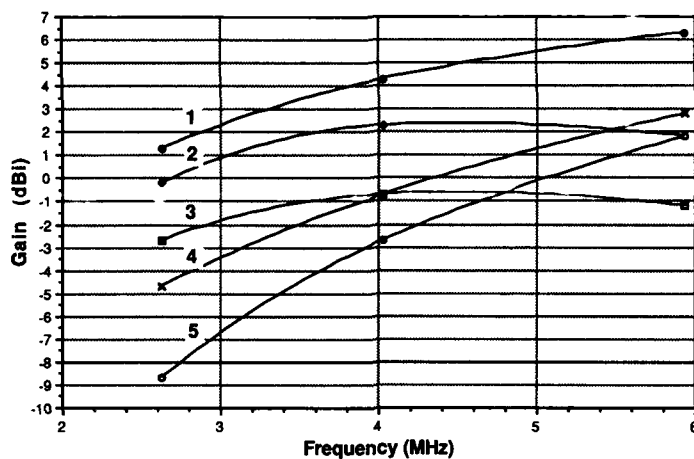


Fig. 22 Measured gain vs frequency for resonant type antennas, i.e. antenna systems that require an antenna system tuning unit (or tuning is a part of the antenna system, as for the mini-loop). The antennas under test are: 1) CRC Prototype Switched Drooping Dipole; 2) CRC Prototype Switched Vee-Dipole (Short Vee); 3) CRC Prototype Switched Vee-Dipole (Long Vee); 4) Rhode & Schwarz Mobile System, Model AK 501 [ref. Stark, 1982]; and 5) Mini-Loop, Racal Model MLA 1E. Ground conductivity of the test site 1 mS/m (estimated), excepting for the Switch Vee-Dipole antennas. Ground conductivity at this test site 10 mS/m (estimated).

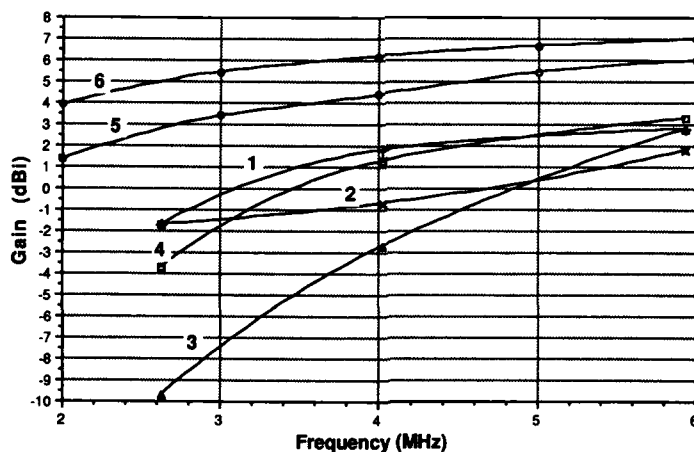


Fig. 23 Measured gain vs frequency for travelling wave dipole type of antennas, Curves 1-4, compared with resonant dipoles at 9.1m (30 feet) and 15.2m (50 feet) Curves 5 and 6. The antennas under test are: 1) AEA Model TWD 215; 2) CSA Model WB DSM; 3) Harris Model RF 1933; and 4) CRC Prototype TWD antenna fabricated and tested by author Petrie. Ground conductivity at the test site was estimated to be 1 mS/m (sandy soil).

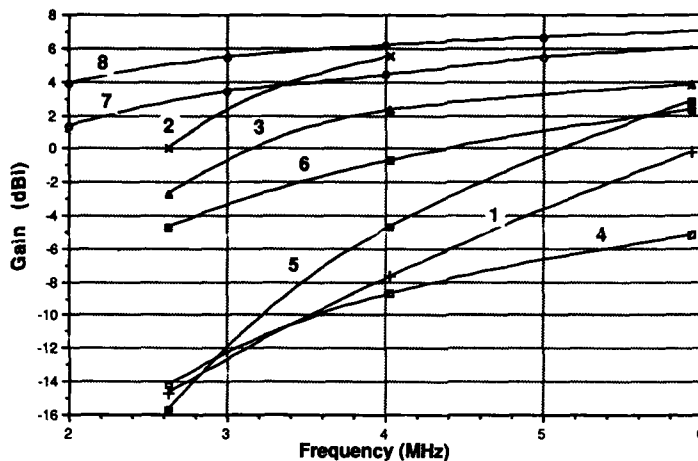


Fig. 24 Measured gain vs frequency for three delta type antennas, Curves 1, 3, and for a miscellaneous set of other types of travelling wave antennas, Curves, 4-6, compared with resonant dipole antennas at 9.1m (30 feet) and 15.2m (50 feet), Curves 7 and 8. The antennas under test are: 1) Small Delta (50 foot mast); 2) Large Delta (100 foot mast), measurements were not possible at the 5.936 MHz frequency because of ionospheric propagation conditions on the day of the test; 3) Tandem Delta, AEA Model TD 230(100 foot mast); 4) Small inverted-V type, B&W 1.8-30; 5) Terminated horizontal folded dipole, B&W 370-15; and 6) Grounded Dipole (dubbed broadband dipole), TCI 613T. Ground conductivity at the test site 1 mS/m (estimated), excepting for the two large (100 foot mast height) delta antennas. Ground conductivity at this test site 10 mS/m (estimated).

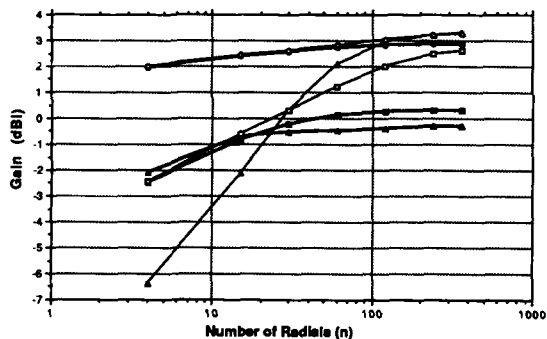


Fig. 25 Theoretical gains (according to NEC-2) for a 3.8 MHz quarter wavelength monopole vs number of buried radials (n), ground conductivity and length of radials. The heavy lines (solid symbols) are for $\lambda/4$ radials; the light lines (open symbols) are for $\lambda/2$ radials. The triangles are for a ground conductivity of 0.15 mS/m; squares are for 3 mS/m; and circles are for 50 mS/m (see table). [after Breakall, 1986].

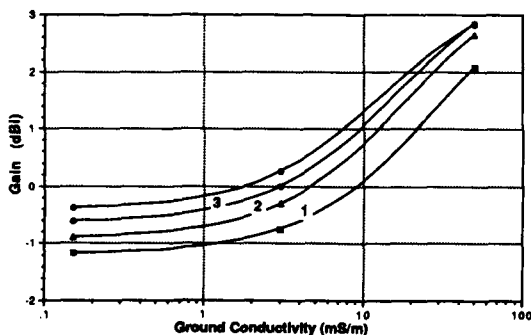


Fig. 26 Calculated gains, according to NEC-2, of a quarter wavelength monopole with four elevated quarter wavelength radials (open symbols) vs ground conductivity and height of the radials above the ground, Curves 1-3; compared with a ground mounted monopole (upper curve with solid symbols) with 120 quarter wavelength buried radials. For curve 1) the height of the radials $h/\lambda = 0.0033$; 2) 0.33; and 3) 0.1.

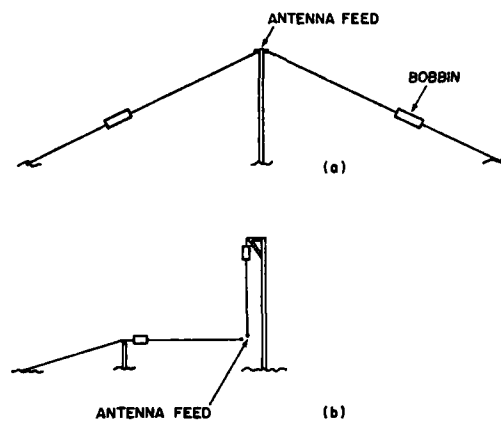


Fig. 27 Sketches of a field portable tactical antenna: a) configured as a conventional drooping dipole; b) configured as a vertical monopole with one radial (dubbed the VE2CV field day special).

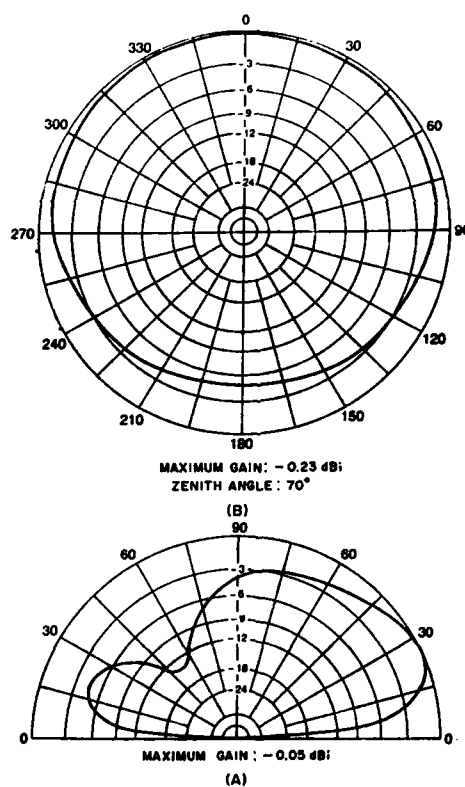


Fig. 28 Azimuthal and elevation pattern for a quarter wave monopole with one radial (VE2CV field day special). Frequency 3.8 MHz, height of radial 15 feet [after Christman, 1988].

ANTENNAS ON LAND VEHICLES FOR IMPROVED COMMUNICATIONS

by
H.K.Lindenmeier
University of the Bundeswehr Munich
8014 Neubiberg, Germany FRG.

SUMMARY

In the meter wave region the time and space dependent superposition of a multitude of waves with great time delay leads to receiving distortions at the location of reception especially in a driven car. This multipath reception, which reduces the quality of FM-reception by noise or by AF-distortions, has been investigated in theory and computer-aided measurements have been made by means of test transmissions from a broadcasting station (100 kW ERP) to evaluate the statistical distribution of the sets of superimposed time delayed waves causing multipath distortion in the receiver. These results are discussed. By means of a specially developed measuring device receiving distortions have also been evaluated statistically for several different types of car antennas under various measuring conditions. The abundance of distortions can be counted for all antennas or for the "auxiliary antennas" in sequence of a distortion on the "main antenna". The result is that all antennas are equivalent as far as their average signal quality is concerned. For application in an antenna-diversity system a distortion detector has been developed which by means of signal processing indicates not only multipath distortions as a result of great time delays between superimposed waves but also indicates noise interference as well as co-channel interference and nonlinear distortions as a result of large signal interference. Main feature of this detector is its extremely short detection time of 25 μ s which is limited to the group delay time of the IF-filters of the receiver. This is a necessity for the application of more than two antennas on the car among which in every moment that antenna is selected which provides the best signal to the receiver. On the basis of this distortion detector in combination with a logic processor the performance of antenna-diversity is investigated as a function of the number of implemented antennas. The results show that the improvement factor increases considerably by the number of antennas applied. For acceptance reasons these antennas are preferably realized as active window antennas. If the improvement factor for a two antenna diversity-system is defined to be q the improvement factor becomes roughly q to the power $N-1$ if N antennas are used. This law has been shown to be true up to 6 antennas on a car. In this respect it is shown that even four antennas can be utilized on the rear window of a car. The relevance of the correlation factor between the distortions of the different antenna signals in comparison with the correlation between the RF-envelopes of the antenna signals is discussed. A computerized measuring equipment for investigation of the system performance is presented. All investigations having been made in the frequency range of FM-broadcasting are also relevant for communications in the general meter wave region.

1. DISTORTIONS CAUSED BY MULTIPATH RECEPTION /1,2,3,4,5,6/

Especially in areas of poor signal supply but also in other regions the reception is often interrupted in form of temporary increase of noise or of distortion of the RF-signal. These distortions result from the space dependent carrier amplitude, which are caused by statistic superposition of a multitude of waves with randomly distributed amplitudes and phases. Therefore due to the motion of the car the amplitude of the received signal varies within a wide level range. This AM-modulation is influenced too by the electrical conducting car body. The time dependent amplitude is displayed in Fig. 1.

1.1 Small differences in delay time

If the time delay between the superimposed waves at the location of reception is less than 1 μ s, the received signal in a standing car is only FM-modulated. In a driven car however an amplitude modulation is produced as shown in Fig. 1. If the RF-amplitude exceeds the noise level of the receiver in spite of a level minimum the receiving system is not disturbed. In times where the RF-amplitude is below the noise level of the receiver noise can be heard. These RF-amplitude minima which are caused by the superposition of a multitude of waves, can not be eliminated by measures at the transmitter side. With a singular antenna system the abundance of distortions can be reduced only by optimization of the receiver sensitivity.

1.2 Great differences in delay time

Further investigations have shown that at different locations noise is present in spite of the fact that the superimposed RF-carrier exceeds the noise level of the receiver. This causes a space dependent level requirement at the receiver input for minimum supply. This effect results from multipath reception with great time delay between the superimposed waves. This is shown in Fig. 1 by the noise level of the receiver and the required level in the case of multipath reception with great time delays. As a result of the space dependent increased minimum required level a temporary receiving distortion exists. Antenna 2 in this example would not be disturbed at any time as a result of the low noise level.

Receiving distortions of this kind result from the superposition of a multitude of waves with great time delays in the range of 1 to 100 μ s. In Fig. 2a the superposition of several reflected waves and a direct wave is shown. The transmission channel between the transmitter and the input terminals of the car radio can be described as a frequency

dependent twoport. The transmission function $S(f)$ of this twoport is equal to a non recursive time delay filter. Fig. 3 shows this structure.

The original spectrum $V_T(f)$ at the input terminals of the transmitter antenna is multiplied with the transmission twoport function and the antenna spectrum $V_R(f)$ at the receiver is obtained from (2) :

$$V_R(f) = S(f) \cdot V_T(f) \quad (1)$$

Generally, $S(f)$ can be described by Eq. (2).

$$S(f) = A_0 + \sum_{i=1}^W A_i \exp(-j\omega\tau_i) \quad (2)$$

In this equation the parameter A_0 represents the transmission factor of the direct received wave and A_i is the general transmission factor of the reflected received i -th wave. τ_i represents the time delay difference between the i -th reflected wave and the direct received wave at the receiver input. The time delay of the direct wave between the transmission antenna and the receiving antenna is irrelevant.

Fig. 4 shows a computer simulated transmission function $S(f)$ with randomly chosen transmission parameters A_i, τ_i . In the case of 2 reflected waves an extreme amplitude and phase response is obtained. With FM-modulation the modulation spectrum covers a wide frequency range and the Bessel-lines are differently weighed and thus the modulation will be distorted.

Fig. 5 explains the modulation distortions of the RF-spectrum. In Fig. 5a the distorted spectrum is displayed for a sinusoidal 1 kHz signal. Fig. 5b shows the spectrum with multipath distortion.

In a driven car the channel parameters vary within a wide range and therefore the amplitude transmission characteristic and the phase transmission characteristic too. This transmission channel can be described by a time dependent dynamic filter.

If the amplitudes of the reflected waves reach the magnitude of the direct received wave and the time delays are in the range of $1 \mu s$ to $100 \mu s$ considerable AF distortions are recognized. Regularly these effects occur often only in a very short period of time. The resulting stochastic distortion pattern generates in the human ear a significant distortion impression. The reason for this is that by application of FM-modulation linear and nonlinear distortions are produced. This is in contrast to AM-modulation, where only linear distortions exist.

1.3. Computer evaluation of distortions caused by multipath reception

In chapter 1.2 it has been shown that multipath reception disturbs the modulation contents. The erroneous signal can be computed as a function of the channel parameter A_i, τ_i and of the transmission modulation. Therefore the resulting voltage $V_R(t)$ obtained from the superposition of a multitude of waves can be calculated by Eq. (3):

$$V_R(t) = V_T(t) \cdot S(t) \quad (3)$$

Eq. (3) is obtained from the fourier transformation of Eq. (1). In the case of FM-modulation the complex transmission signal can be described by

$$V_T(t) = |V_T(t)| \cdot \exp(-j\omega_T t + j\phi(t)) \quad (4)$$

with $\phi(t)$ representing the modulation phase angle. Eq. (3) applied to Eq. (4), delivers the voltage $V_R(t)$ at the receiver input terminals.

$$V_R(t) = A_0 \cdot |V_T(t)| \cdot \exp(-j\omega_T t + j\phi(t)) + \sum_{i=1}^W A_i \cdot |V_T(t)| \cdot \exp[-j\omega_T(t-\tau_i) + j\phi(t-\tau_i)] \quad (5)$$

The computation of the modulation phase $\phi(t)$ of $V_R(t)$ and the following differentiation of $\phi(t)$ gives the instantaneous frequency deviation $\Omega(t)$.

$$\phi(t) = \arctan \frac{\text{Im}(V_R(t))}{\text{Re}(V_R(t))} = \arctan \frac{A_0 \sin\phi(t) + \sum_{i=1}^W A_i \sin(-\omega\tau_i + \phi(t-\tau_i))}{A_0 \cos\phi(t) + \sum_{i=1}^W A_i \cos(-\omega\tau_i + \phi(t-\tau_i))} \quad (6)$$

$$\Omega(t) = \dot{\phi}(t) + \frac{\sum_{i=1}^W [(\phi(t-\tau_i) - \phi(t)) (A_0 A_i \cos \gamma_i(t) + \sigma_i(t))]}{A_0^2 + \sum_{i=1}^W (A_i^2 + A_0 A_i \cos \gamma_i(t) + \sigma_i(t))} \quad (7)$$

with

$$\gamma_i(t) = -\omega_T \tau_i + \phi(t - \tau_i) - \phi(t),$$

$$\sigma_i(t) = 2 \cdot \sum_{j=i+1}^W A_i A_j \cos(\omega_T(\tau_j - \tau_i) + \phi(t - \tau_i) - (t - \tau_j)).$$

The difference between the receiving frequency deviation of the received signal and the frequency deviation of the original signal is the distortion frequency deviation $\hat{\beta}(t)$:

$$\hat{\beta}(t) = \frac{\sum_{i=1}^W [(\phi(t - \tau_i) - \phi(t)) (A_0 A_i \cos \gamma_i(t) + \sigma_i(t))]}{A_0^2 + \sum_{i=1}^W (A_i^2 + A_0 A_i \cos \gamma_i(t) + \sigma_i(t))} \quad (8)$$

Since A_i can not be assumed to be small for any i all terms in Eq. (8) have to be considered. An analytic evaluation of Eq. (8) is not possible, therefore a numerical computer method was applied to investigate the distortion factor.

The distortion factor has been evaluated by computer simulations for a sinusoidal AF-signal of 1 kHz and two reflected waves.

Fig. 6 shows the distorted output signal of an ideal FM-demodulator. The distortion spikes are evident. The resulting distortion spectrum at the FM-demodulator output is distributed within a very wide frequency range. In the simple case of only one reflected wave with an amplitude of 0.5 normalized to the directly received wave a distortion factor as in Fig. 7 is obtained at the output of an ideal FM-demodulator. The distortion is displayed as a function of the time delay between the 2 waves and the AF-frequency. It can be seen that even short time delays below $10 \mu s$ cause severe distortions. Fig. 7a presents the dependence on the related amplitude of the echo wave.

1.4 FM-AM conversion with multipath reception

Since multipath reception is always correlated with FM-AM conversion this AM-modulation can be advantageously used for detection of multipath reception. But this detection method has no high recognition security because in a mobile receiving system the receiving level varies violently with the variation of the receiving situation.

In the following, this AM-modulation will be computed. The absolute value of $\underline{V}_R(t)$ describes the AM-modulation:

$$|\underline{V}_R(t)| = \sqrt{\text{Re}\{\underline{V}_R(t)\}^2 + \text{Im}\{\underline{V}_R(t)\}^2} = \sqrt{A_0^2 + \sum_{j=i+1}^W (A_i^2 + A_0 A_i \cos \gamma_i(t) + \sigma_i(t))}. \quad (9)$$

Fig. 8 shows the computed results for parameters identical with those of Fig. 4.

If the results of Fig. 6a are compared to those of Fig. 8 the distortion spikes are correlated with the minima of the RF-amp'itude. This phenomenon can be easily explained with the complex vector diagram. In this case two waves of different amplitude and phase are considered. Therefore each of both complex vectors move with different angular speed if a modulation signal exists. In Fig. 9 only the relative movement of the complex vector of wave 1 in relation to wave 0 is considered.

Assuming a great frequency deviation and time delays in the range of 10 - 100 μs the phase angle between both waves can reach the factor 2 or more. With a periodic modulation signal several complete turns of the complex vector of wave 1 related to wave 0 will occur.

The resulting complex vector will get several minima. The instantaneous distortion frequency deviation is represented by the differentiation of $\psi(t)$. Any sufficiently great difference in angular velocity between both complex vectors generates distortion spikes. Therefore with the example of a constant difference in angular velocity the differential

of the resulting complex vector of the waves is increased in the range where both waves are of opposite phase. If the absolute values of both complex vectors are equal, spikes of infinite height are obtained. A similar phenomenon exists with superposition of more than two waves.

1.5 Response of the ratio detector in case of multipath reception

An ideal FM-demodulator evaluates only the instantaneous frequency deviation of the input signal. This means that even under the assumption of a total suppression of the RF-carrier fluctuation the FM-demodulator operates correctly. An ideal FM-demodulator has a linear and an infinitely wide demodulation characteristic, which is not true for a real FM-demodulator. Therefore the demodulated signal will be more distorted in comparison to that with an ideal demodulator.

Fig. 10 shows the demodulation characteristic of a ratio detector under the assumption of a sufficiently high signal level. The demodulation characteristic is only linear in the very vicinity of the f_f -frequency. With greater frequency deviation limiting effects occur and in a further step a reciprocal weighing of the frequency deviation occurs. If the receiver input signal is undistorted the frequency deviation covers only the 5-th part of the linear demodulation characteristic between minimum and maximum and therefore a FM-demodulation distortion does not exist. If the receiver input signal consists of superimposed waves with great time delays between the waves, the frequency deviation can exceed more than 10 times the frequency deviation of a normal signal. These frequency deviation spikes will be limited and by that the distortion factor will be increased.

Generally, it is assumed that the ratio-detector suppresses the amplitude modulation. This is only true for a small amplitude modulation factor. In the case of multipath reception a high amplitude modulation factor is obtained (see Fig. 8). If the RF-level remains under a certain threshold value V_d , the AF-signal is a function of the RF-level (Fig. 11), because the distortion spikes are correlated with a minimum RF-level. This effect causes an increasing distortion factor.

In Fig. 12 the distortion factor is plotted versus the time delay between a direct and one reflected wave and versus the AF-frequency of a sinusoidal signal by using a ratio detector. The peak frequency deviation is 75 kHz and the RF-level of the related reflected wave is 0.9. If the distortion factor, which is obtained from the ideal FM-demodulator (see Fig. 12a) is compared to the distortion factor of the ratio detector (see Fig. 12b), the distortion factor of the ratio detector is greater especially at low AF-frequencies and time delays. If the peak frequency deviation and the RF-level of the reflected wave related to the direct received wave decrease, the difference between the distortion factor obtained from the ratio detector and the ideal FM-demodulator is smaller and will become equal.

The pulse-counter frequency discriminator represents a good approximation for an ideal FM-demodulator. The demodulation characteristic of this kind of discriminator is only limited by the supply voltage. If the output signals of this discriminator and the regular ratio detector are compared no difference can be heard. The reason for this is found in the fact that the intermediate frequency deviation is app. 20 kHz, so that nonlinear distortion effects of the ratio detector are negligible.

The investigations of several FM-demodulators show that the application even of an ideal FM-demodulator would in no way solve the problems of multipath reception distortion.

2. SUPPRESSION OF ECHO DISTORTIONS UTILIZING A VARIETY OF ANTENNAS

From the curves in Fig. 1 it can be learned that the minima of RF-amplitude with different antennas on a car regularly do not occur at the same time. In the past many attempts have been made to make use of this effect by selecting the receiving signal from that antenna which momentarily provides the best signal. In lack of a sufficiently fast operating distortion detector no antenna diversity system has been presented in the past which could make use of the advantage of two and more antennas on a fast driven car. In ch. 2.3 of the present paper an antenna diversity system implementing a fast indicating distortion detector is described. In the following the efficiency of such a system under the assumption of such a detector is investigated as a function of the number of antennas applied.

The results show that the improvement factor is considerably increased by each additional auxiliary antenna. For this reason a multitude of up to 12 antennas mounted on a car have been measured. The installation of these antennas is shown in Fig. 13 where four antennas (1, 2, 5 and 6) are mounted on the fenders, further six antennas (7 to 12) are mounted at the shown locations on the roof and an additional pair of windshield antennas (3 and 4) is fixed on the front window. It is well known that the correlation factor increases with decreasing distance between adjacent antennas. In order to find the improvement factor of distortion suppression also as a function of different antenna locations on a car the set of antennas presented have been measured. For this reason a special equipment had to be developed.

2.1 Computer aided antenna measurements

In order to obtain objective measuring results a radiated test signal was necessary since multipath distortions are a function of the modulation. For this reason in cooperation with the Bavarian Broadcasting Corporation a test signal with 1 kHz modulation and a maximum frequency deviation of 35 kHz has been radiated during the night with a power of 100 KW(erp) from several stations along the north rim of the Bavarian alps. Since the farther north of the alps being a flat area, and with Munich being a densely built up area, and the north rim of the alps being mountainous all main different receiving conditions which are important in practice are available. Therefore studies could be made in the flat areas with low multipath distortions, and in high multipath city areas with small delay times between the superimposed waves, and in high multipath areas in the mountains with great delay times between superimposed waves as well.

The measuring device in the car exists of a scanner, which sequentially switches the antennas to the receiver for one period of the audio frequency cycle each (see Fig. 14). The receiver contains a broadband and highly linear FM-demodulator and an AM-demodulator of a wide dynamic range for amplitude detection. The FM-detected signal is filtered by a lowpass of 500 kHz and a trap of 1 kHz of approximately 100 Hz bandwidth and this signal is integrated by an integrator. The instantaneous amplitude is also integrated and by means of a sample and hold circuit is available at the same time at the input of the selector. This alternately reads the FM- and the AM-values into the system voltmeter digitizing the values and providing them to the system computer of type HP 9816 S. These values are stored on the floppy disks and statistically evaluated in the laboratory after the drive. After every scanning cycle the instantaneous error of the receiving signal of each antenna is available as a stored value in the computer.

2.2 Available reduction of echo distortions by antenna diversity

As a basis for a realistic judgement of the distortion the definition of a subjectively acceptable upper limit of the audible distortion impression for a reference is necessary, and the system delivers a digital information of the presence of a distortion by comparing the instantaneous value with this reference. The corresponding electrical value of this reference can therefore serve for an evaluation of a well defined 'distortion degree', which describes the likelihood of the presence of a distortion with only one antenna related to the likelihood of a distortion with the antenna diversity system to be regarded. The ratio of these likelihood values can therefore be regarded as the improvement factor q , which reasonably is expressed logarithmically as:

$$q_{dB} = 20 \log q. \quad (10)$$

This factor is relevant to express the audible factor of improvement with reference to the receiving condition, which is obtained with one antenna in the receiving situation regarded.

Fig. 15 shows the improvement factor q_{dB} versus the number N_a of auxiliary antennas applied. The curve described by the q_{dB} -values for the different values of N_a is found to be almost a straight line. A rough approximation of this curve is given by:

$$q_{dB} = 20N_a - N_a^2 \text{ with } 0 \leq N_a \leq 3 \quad (11)$$

This diagramm has been measured in the City area of Munich with a radio signal being radiated from a weak station 80 km away from Munich. The most interesting aspect is found in the fact that the improvement factor increase with a median value of roughly 17 dB per additional auxiliary antenna. This shows that the investigation of the efficiency of antenna diversity may not be reduced to only one auxiliary antenna. Antenna 1 served as main antenna. Auxiliary antennas were antenna 2, 5 and 6 for $N_a=1,2$ and 3 respectively (s. Fig. 13).

With the measurements in Fig. 16, which have also been made in city areas the maximum number of auxiliary antennas was chosen even higher, and up to $N_a=4$ has been evaluated. The results show a curve, which with increasing N_a ascends however the increase of the q_{dB} -factor decreases with increasing N_a of auxiliary antennas applied. For a rough estimation of the available improvement factor the following formula is valid:

$$q_{dB} = 20N_a - 2.5N_a^2 \text{ with } 0 \leq N_a \leq 4 \quad (12)$$

In this context it should be stated that the available improvement factor is a function of the signal quality being received with the normal (single antenna) system. This signal quality can be defined as the percentage of time where the signal is received undistorted.

In city areas the time delay between superimposed waves is comparatively small and the number of waves of different time delays contributing to multipath distortions regularly is comparatively low. This is in contrast to the situation in mountain areas where the median value of the difference in delay time between the superimposed waves is in the order of magnitude of 40 μ s corresponding to a detour of 12 km in propagation. Due to this fact the average multipath distortions are much more severe in mountain areas. In Fig. 17 the improvement factor measured in a heavily multipath distorted region is plotted versus the various numbers N_a of auxiliary antennas. The measurements have been

made around the lake of Tegernsee. The curve obtained is almost a straight line indicating the law of equal increase of the logarithmic improvement factor per additional auxiliary antenna. Even though there is only little space between adjacent antennas the cross correlation factor between the output voltages is still low enough in order to get the measured increase. In the case regarded the signal supply is very poor and therefore the signal quality, which is the percentage of time in which the reception is undistorted, is low. The curve can be approximated by the following law:

$$q_{dB} = 7N_a \quad \text{with } 0 \leq N_a \leq 6 \quad (13)$$

As a result of the poor signal supply the increase of the improvement factor per additional auxiliary antenna is less than in city areas. In a mountain valley where good signal supply was found the measurements in Fig. 18 have been made. In this case the improvement factor versus N_a can be approximated by:

$$q_{dB} = 21N_a - 2.5N_a^2 \quad \text{with } 0 \leq N_a \leq 3 \quad (14)$$

With low numbers of N_a the increase of improvement factor is roughly 20 dB per additional antenna and is found to be as large as in city areas. This is true in spite of the large differences in delay time between superimposed waves. As a result of the good signal quality however the further increase of the improvement factor per additional auxiliary antenna is less if N_a is a great number.

For a better survey all curves measured in city areas are plotted in one diagramm and the dotted range indicates the region wherein the improvement factor can be expected with a given number N_a of auxiliary antennas applied (Fig. 19). A similar diagramm is shown in Fig. 18 for the reception in mountain areas. Finally all curves measured in city- and mountain areas as well are shown in Fig. 19 and the dotted range indicates the improvement factor to be expected in a large variety of receiving areas of different characters.

2.3 Experiments with antenna diversity with a multitude of antennas

In the past antenna diversity systems have been investigated mainly for two antennas on the car. Developments having been published in the past make use of two receiver tuners at the input of the diversity system. Main problem with such a system is the detection time for realisation of a signal distortion due to delayed superimposed waves, noise, and co-, and neighbour- channel interference as well. In the present paper the receiving results are described, which are obtained by means of an extremely fast indicating distortion detector. Due to the detection time of only 25 μ s it is possible to control an antenna selector which sequentially chooses the best antenna among a variety of N antennas mounted on a car. As a further result of the short detection time only one single tuner at the front end of the receiver is necessary. Fig. 20 shows a principle diagram of the antenna diversity system under test. Antenna 1 is defined to be the main antenna for both AM- and FM-reception. Antennas 2 to N are auxiliary antennas for FM-reception only with $N_a = N-1$. The antenna selector switches by means of switching diodes from one antenna input to the next if the distortion detector within the antenna selector indicates a distortion of the if-signal supplied from the FM-radio. The short detection time also allows the simultaneous application of more than two antennas among which the instantaneous receiving antenna is selected. A detailed description of the diversity processor is given in chapter 3.

By means of this system many tests have been made in a vast variety of different receiving areas and receiving conditions and with different numbers N_a of auxiliary antennas applied. These tests have been made both subjectively by listening and judging the quality of the obtained signal, and by objective measurements and statistical evaluation. In city areas where the signal supply regularly is comparatively good and the percentage of time in which distortions are audible with a single antenna system is low only two antennas ($N_a=1$) are sufficient to reduce the distortions to a tolerable level. In heavily distorted areas however, such as mountain areas or areas with poor signal supply, more than two antennas should be used on the car. Good results have been obtained with a variety of window antennas.

In order to get objective results the measuring system described in Fig. 14 has been extended to the system in Fig. 21 and the signal provided by the output of the antenna selector of the diversity system via channel $N+1$ in Fig. 21 is compared to the different signals of the N antennas on the car. By means of this measuring set the range in Fig. 19 has been verified.

2.4. Distortion reduction by an adaptive equalizer

The transmission twoport in Fig. 2b describing the transfer function between transmitting antenna and the car antenna gives rise to the idea to compensate for the phase and amplitude errors by means of an adaptive equalizer to be implemented in the FM-radio. An equalizer of this kind has been developed. It corrects the total frequency response from transmitter-aerial to receiver if-output and removes echos and the related distortions from the signal. The equalizer has an adaptation time of about 4 milliseconds and can handle echos up to 70 microseconds delay time. The described diversity-system together with this equalizer promises the best possible reception under multipath conditions in moving cars: no audible interruptions, distortion or other discontinuities. It has been shown however that an equalizer of this kind can process no more than two waves at a time.

2.5. Registration of disturbing echo-waves with $\tau_1 > 10 \mu s$

For the detailed design of the diversity processor in Fig. 20, statistics of the nature of the FM-radio distortions are necessary. For this reason the test signal in Fig. 22 was radiated from an FM-radio station in the Bavarian Alps ($P=100kW$ erp; height: 1850m) and the number and the amplitudes of the measured sets of echo-waves, and the differences in propagation time were evaluated by means of the pulses within the test signal. The distortions being generated by the specific set of waves have been found by means of the sinusoidal part of this signal. With a maximum frequency deviation during the pulses of 75kHz the probability distribution of the number of waves forming a set of waves that generates an audio distortion of at least 20% is presented in Fig. 23. The appertaining probability distribution of the delay time is shown in Fig. 24, that of the wave amplitudes belonging to a set related to the total amplitude of the resulting carrier in Fig. 25, and that of the distortion factor in the audio signal in Fig. 26. Measurements have been made in mountain areas (shaded) and in flat areas as well. Figures 23, 24 and 26 are of special interest, coming to the conclusion that in average the number of echo waves and the difference in delay time is greater in mountain areas than in flat areas. This results in a considerable greater average of distortion factors in mountain areas. Former investigations have shown that the application of signal processing techniques in the receiver for correction of erroneous signals might only be realized for a situation with only one echo wave. From Fig. 23 it is learned that antenna diversity must be considered the only means to reduce distortions in the receiver.

The measurements show that the potential of an antenna diversity system increases considerably with the number of antennas supplied. The deteriorating influence of cross correlation between antenna output signals as a result of narrow spacing in practice is less than it could be expected from former theoretical studies. With low numbers of auxiliary antennas $N_a = 2$ in all cases an equal increase of the logarithmic improvement factor per additional antenna has been found. In regions with great differences in delay time between superimposed waves in combination with poor signal quality the improvement factor obtained with a certain number of auxiliary antennas is less. Therefore the potential of antenna diversity with a low number of auxiliary antennas in urban areas is greater than in typical mountain areas. This measured potential of improvement has been verified with a realized antenna diversity system applying a specially developed multipath distortion detector with error indication time of only 25 μsec in combination with a Philips car radio of type MCC 914. Practical experiments with a variety of window antennas, where two antennas were built into the windshield and two antennas printed onto the rear window, delivered similar results. For a limited number $N_a = 1$ the further improvement of the reception by means of an adaptive equalizer is advantageous.

3. DIVERSITY PROCESSOR FOR ALL KINDS OF DISTORTIONS /6,7,8,9/

3.1. Distortions by noise and by other transmitting stations

As a result of multipath propagation of the rf-carrier wave distinct amplitude minima of the received signal occur during mobile radio reception. As in Fig. 27 the level variation in the region of such a minimum is often in the range of 30 - 40 dB and different kinds of interference such as noise-, co-channel-, adjacent channel-, and intermodulation interference may occur during those time periods wherein the desired signal level does not exceed the interference level. The superposition of waves with a time delay of $>10 \mu s$ as discussed above also leads to audible distortions if an amplitude condition is satisfied, which is marked by the shaded range in Fig. 27. Therefore it is obvious that the distortions are strongly correlated with amplitude minima and the number of minima per distance and their depths are a measure for the signal quality. In a Rayleigh field the signal received with an omnidirectional antenna is presented by the uninterrupted curve in Fig. 28 versus the driven distance. The interrupted curve is obtained if a directional antenna pattern is applied generating less deep signal minima and therefore causing less distortions. Since it is obvious that a directional antenna cannot satisfy the requirements for a mobile receiving antenna the receiving problem can only be solved by means of a time variable antenna system.

If the signal level does not exceed the interference level forming the threshold in Fig. 29, where the amplitudes of two antenna signals on a car are plotted versus the driven distance, a distortion occurs. This is indicated by the binary signals below. With p_1 and p_2 representing the fraction of time for a binary distortion in signal 1 and signal 2 respectively the likelihood for the simultaneous occurrence of a distortion in both signals is nearly $p_1 \cdot p_2$. Assuming an extremely fast operating distortion indicator a scanning diversity system as in Fig. 20 may be applied, making use of a multitude of receiving antennas on a car with only one receiver.

Measurements have shown that with increasing number N of antennas applied the fraction P_d of the time in which the threshold of distortions plotted in Fig. 34 is exceeded is reduced exponentially and that this improvement is increased with the signal quality Q_s . This can be shown from the following assessment, which in a first approach is made under the assumption of equal distortion probabilities

$$P_1 = P_2 = \dots P_N = P \quad (15)$$

and negligible correlation of the distortions found at the antennas 1...N.

$$Q_s = 20 \log (1/P) \quad (16)$$

may be considered the signal quality with the single antenna system. Hence the probability of a distortion with the N-antenna diversity system with $N_a = N-1$ auxiliary antennas is found to be the joint probability

$$P_d = P_N \quad (17)$$

If the improvement factor q is defined as P/P_d its logarithmic value

$$q_{dB} = 20 \cdot \log q \quad \text{reads as } q_{dB} = N_a \cdot Q_s. \quad (18)$$

The shaded range in Fig. 19 represents measured curves of the improvement factor versus the number N_a of auxiliary antennas under various receiving conditions. With higher numbers of N_a the correlation between the antenna signals can not be neglected, which causes the curves to bend. In contrast to a selection diversity system the described scanning diversity system can take advantage from a great number N of antennas in practice.

3.2. Diversity processor

In /8/ it is shown that all interferences of the various kinds discussed in ch.3.1 go along with distinct erroneous pulses of the instantaneous frequency deviation of the frequency modulated rf-carrier in combination with sharp erroneous pulses modulating the envelope of the carrier forming deep minima as displayed in Fig. 6a and Fig. 8. The coherence between the maxima of the frequency deviation pulses and the amplitude minima serves as an excellent extremely fast evaluable criterion for distortion recognition in the diversity processor shown in Fig. 30.

The block diagram of the diversity processor is displayed in Fig. 30. The if-signal representing the rf-signal behind the mixer of the receiver is demodulated in amplitude and frequency in different branches of the circuit. After demodulation the signals are shaped in steepness by means of nonlinear circuits. In this way the significant erroneous pulses in frequency deviation and rf-envelope curve are pronounced. Both signals are lead to a comparator circuit the outputs of which are connected to the input of a binary logic. The logic AND-circuit indicates the simultaneous presence of a pulse in the frequency deviation and the rf-envelope. If a distortion pulse is recognized in the frequency deviation and the envelope of the if-carrier at the same time the binary logic moves the signal selector without delay time into a different position searching for a less distorted antenna signal. By means of an appropriate feedback of the sequence of antenna switching to the thresholds of the two comparators a dynamic adaptation of the diversity processor to the receiving situation can be obtained. This means that in areas with high average distortions the processor reacts less sensitive in a way that the movement of the antenna selector does not exceed a tolerable sequence.

4. ANTENNAS FOR ANTENNA DIVERSITY

4.1. Minimum required signal level with a single antenna

Up to now no minimum required signal level for mobile FM broadcast reception has been defined. Due to the random character of the received signal in Fig. 35 its relevant level can only be determined by means of statistical characteristics such as the exceeding probability distribution of the instantaneous logarithmic signal amplitude at the input of the carradio (see Fig. 36). This curve can be characterized by the median voltage level V_m , which is exceeded in 50% of the time. Normally the steepness of the slope is characterized by the level difference "b" between the 10 % and the 90% points, which in a Rayleigh field is found to be app. 13 dB. A working group of the German Society of Car Producers has defined minimum required signal levels under various receiving conditions. $V_m = 24 \text{ dB } \mu\text{V}$ has been found for FM reception in urban areas and $V_m = 13 \text{ dB } \mu\text{V}$ in non urban and flat areas of reception. These low signal level requirements are only valid in areas where no multipath distortions due to great differences in delay time ($\tau = 3 \mu\text{s}$) are observed /10/.

4.2. Antenna characteristics

Very often the main specification relevant to describe the performance of an FM car antenna is considered to be the directional diagram as to be seen in Fig. 31. However the artificial situation of only one wave being incident at the location of reception, which with pattern measurements is assumed, does not meet the physical conditions of FM carradio reception in practice. In fact at any location of reception there is a multitude of coherent waves of complex amplitudes $A_1 \dots A_w$ incident from various directions 1...w in azimuth and elevation. With $C(\varphi_v, \theta_v)$ representing the complex antenna factor and $A_v \cdot C(\varphi_v, \theta_v)$ representing the contribution of wave w to the complex voltage amplitude at the receiver input it becomes obvious that the total complex amplitude at the receiver in a standing car reads as:

$$V = \sum_{v=1}^w A_v \cdot C(\varphi_v, \theta_v) \quad (19)$$

and can cancel out even if the diagram would be omnidirectional ($C(\varphi, \theta) = 1$). In a driven car the set of complex wave amplitudes varies very rapidly versus the time in magnitude, phase, angle of incidence and number of waves to be considered, so that the receiver

input voltage reads as:

$$V(t) = \sum_{v=1}^W A_v \cdot C(\varphi_v(t), \theta_v(t)), \quad (20)$$

and varies within a wide range as to be seen in Fig. 35. The validity of these considerations can be seen from Fig. 32, where the amplitudes of the instantaneous set of waves are plotted versus their angles of incidence. These angles have been found by measuring the doppler frequency shift f_D of an unmodulated and highly frequency stabilized FM broadcasting carrier which for this purpose has been radiated by the Bavarian Broadcasting Corporation with a power of 100kW erp. f_D is found from:

$$f_D = f_C \cdot v/c_0 \cdot \cos \varphi; \quad f_C: \text{Carrier frequ.}; \quad v: \text{car speed};$$

c_0 : velocity of light.

With $V(t)$ being of random character it is obvious that only little can be predicted with respect to a definition of a desirable antenna factor representing the directional diagram. In contrast hereto the exceeding probability distribution has been found to be a reliable and an easy to measure means to characterize the antenna performance.

4.3. Active windshield antenna

Windshield antennas are often blamed of being "more directive". Fig. 31 shows that there is no criterion that the diagram of antenna B should be preferred to that of antenna A. In fact the exceeding probability distribution with an efficient windshield antenna is of the same profile as obtained with a rod antenna. Fig. 33 presents an active FM-AM-antenna which is implemented in the windshield in the packaged glass. The design is optimized with respect to the reception of horizontally, vertically, and circularly polarized FM-waves and is adapted to the individual car bodies by appropriate choice of d and l of the conductor. The exceeding probability distribution of the received antenna signal is compared to the whip antenna in Fig. 36. It shows that both characteristics are equivalent.

As discussed in ch.1 multipath reception disturbs the modulation and harmonics are generated as a function of the difference of propagation time τ between a wave of amplitude A_0 and an echowave of amplitude A_1 if only two waves are existent. Fig. 34 shows the distortion level versus the driven distance measured with this detector and the antennas on the car in Fig. 33. The median distortion level with the windshield antenna was app. 10% less than with the rod. Obviously the peaks of distortions exceeding an audible threshold with the two antennas are only very little correlated. Therefore antenna diversity can provide a substantial improvement of signal quality to the system.

5. PROFESSIONAL MULTIPLE ANTENNA DIVERSITY SYSTEM /11/

5.1. Scanning diversity processor utilizing a linear combination of signals

Fig. 41 shows a diversity system consisting of an antenna system, an FM-receiver and an external diversity processor. The basic principle of this processor is described in ch.3. At the input of the diversity processor linear combinations of the antenna signals $A_0 \dots A_N$ may be derived and connected to the inputs $EO \dots EN$ of the electronic switch. Obviously an arbitrary number of inputs of the switch can be created. This would allow to operate the system with a comparatively small number of antennas. Linear combinations of the signals however are not as efficient as additional antennas even though these antennas often must be closely packed on the vehicle. For the investigation of this fact measurements in urban areas and in mountain areas have been made. In the following the performance of linear combinations of antenna signals are investigated for an example of 3 antennas. With this example linear combinations of antenna signals are derived forming the sum and the difference by means of the matrix in Fig. 41. Figs. 42a and 42b present the available improvement in signal quality. With 3 antenna signals 3 further signals are generated by linear combinations. The results show that with multipath distortions due to great differences in delay time between the superimposed waves the gain in signal quality with an additional linear combination is greater than in urban areas where this kind of distortion is less likely. In urban areas the distortions produced by other transmitters such as co- and adjacent channel interference and intermodulation distortions are dominant. For this kind of distortions linear combinations do not help to improve the signal quality.

5.2. Two antennas in the windshield and two on the rear window

The antenna arrangement is displayed in Fig. 39. Antenna 1 is equivalent to that in Fig. 33. Antenna 2 and 3 are derived from the subdivided heater structure on the rear window. Antenna 4 is a quarter wavelength wire in parallel to the right window post. In Fig. 40 the statistically evaluated levels of the rf-amplitude and the distortion levels are displayed. The exceeding probability of the signal levels show that the likelihood of small signal levels representing signal fadings at the receiver input is very small with the diversity system. The appertaining distortion levels show that with 4 antennas the likelihood of distortions is much less than with the single antenna system.

5.3 Integrated 4-antenna system on the window of a car

The most amazing result of the following investigations is the fact that diversity

antennas may be closely packed with very narrow space between the samples. Most theories on diversity systems claim for a certain distance between the antennas applied in order to obtain the necessary decorrelation between the antenna signals. It has been shown that making use of the resonances of a car body and utilizing a variety of different antennas the desired decorrelation can be realized even though all antennas are placed on only one of the car windows. Fig. 44 shows 4 FM-antennas two of which are derived from the heater structure and two further are placed in the remaining gaps between the heater and the window frame. This system is compared to a car utilizing 4 standard mast antennas of 1m length. By means of the test equipment in Fig. 49 the signal quality and the correlation of signals and the distortions have been evaluated by the test computer. The plot in Fig. 47 proves that in spite of the compact antenna system in Fig. 44 the correlation factors of the distortions in the signals are small enough to obtain a diversity performance which is practically as good as with theoretically uncorrelated distortions. Very often the correlation of the amplitudes of the antenna signals is regarded to predict the efficiency of a diversity system. All investigations however show that this correlation factor is not an appropriate criterion to predict the efficiency of the system. Since with FM-communication the amplitude of the rf-signal gives no direct indication of the signal quality the distortion is the only relevant criterion. It is therefore quite natural that the correlation factors of the measured signals and the measured distortions are different. The fact that the correlation factors of the distortions are less than those of the signals can be advantageously used to install a multitude of antennas on a comparatively small space on a vehicle.

With all antennas being located in the same region of the car an external diversity processor such as in Fig. 41 is advisable. In cases where the antennas are spaced apart as far as possible the diversity processor should be installed inside the receiver set as the example in Fig. 43 shows.

To give an impression of the contribution of an additional antenna to the distortion reduction Figs. 48a...48d may be regarded. The measurements have been taken during the same drive with only one antenna in Fig. 48a, 2 antennas in 48b, 3 antennas in 48c and 4 antennas in 48d. In this case 4 antennas were enough that at no location the threshold of audibility has been exceeded by the distortions.

6. MEASURING EQUIPMENT

In order to investigate the performance of an antenna system the measuring equipment displayed in Fig. 49 has been developed. This system allows the simultaneous measurement of 8 antennas which in a fast sequence are connected to the input of an FM-receiver by a clock. This clock is produced by the system clock of a digital voltmeter. The received signal is transformed by two mixing stages and its amplitude is controlled constant in amplitude before the signal is demodulated in amplitude and frequency. The gain is controlled by means of the PI-controller from the AM-signal. The rf-level is indicated by the control voltage of the PI-controller. In this way the RF-level is measured from 10 dB μ V to 80 dB μ V with an accuracy of ± 1 dB. The MPX-signal at the output of the FM-demodulator is filtered via a highpass filter of 60kHz cut off frequency the output signal of which indicates the distortion of the rf-signal. With this procedure the distortion of the antenna signals can be measured in a normal broadcasting signal. In the time of 10ms 8 antennas can be registered as far as their signal levels and signal distortions are concerned. With a driving speed of 15km/h the driven distance during a complete cycle of measurements is not greater than 4cm. After each cycle of measurements the system computer needs a time of 200msec for data storage. Approximately 4000 cycles of measurements may be registered during one drive and stored in the computer. At the end of the drive the stored data can be either immediately processed or stored on the mass storage such as the floppy disk. The following characteristics can be processed from these data:

- Median values of the rf-level and the distortion level
- Correlation between the signals of different antennas concerning level and distortion respectively
- Display of level and distortion versus time for each antenna signal
- Exceeding probability distribution of the rf-level and the distortion
- Plots of directional diagram of all antenna signals for circular drives
- Computational evaluation of the efficiency of a multiple antenna diversity with regard to minimum distortion of the diversity system utilizing the antennas under test.

7. COMPUTATION OF CAR ANTENNA CHARACTERISTICS

For a more efficient development of antennas on a vehicle computational methods for investigation of antenna characteristics are inevitable. The antennas applied to the vehicles in this project have been computed from a wire grid model as presented in Fig. 50a wherein for example an antenna in the windscreen as displayed in Fig. 33 is placed. The computations are made utilizing the well known program NEC. In Figs. 50b and c the three-dimensional radiation characteristic is plotted for horizontal polarization and vertical polarization respectively. In spite of the high radiation density for steep radiation angles of this antenna it has as good characteristics in the horizontal plane as the standard rod antenna.

REFERENCES

- /1/ Tschimpke, L., "Space Diversity for Mobile Reception of Meter Waves in Urban Areas", University of the Bundeswehr Munich, FRG, Doctor Thesis 1981.
- /2/ Jakes, W. C., "Microwave Mobile Communications", John Wiley & Sons, New York 1974.
- /3/ Lee, W. C. Y., "Mobile Communications Engineering", McGraw-Hill Book Company, New York 1982.
- /4/ Lindenmeier, H. and Reiter, L., "Distortions in FM Car Radios caused by Multipath Reception and their Reduction by means of Antenna Diversity", Proc. Ursi Commission F 1983 Symposium, Louvain, Belgium, June 1983 (ESA SP-194), p. 37 - 44.
- /5/ Lindenmeier, H. K. and Manner, E. J., Sessink, F., "FM Antenna Diversity Experiments in Philips Car Radios in Theory and Practice", Society of Automotive Engineers, Detroit 1986.
- /6/ Manner, E. J., "Distortions and their Suppression by means of Multiple Antenna Diversity for FM Radio Reception in Cars", University of the Bundeswehr Munich, FRG, Doctor Thesis 1985.
- /7/ Lindenmeier, H. K., "Advances in FM Carradio Reception", Proceedings of the Int. Soc. of Antennas and Propagation of Japan, Conference Proceedings 1985, p. 437 - 440.
- /8/ Lindenmeier, H. K. and Manner, E. J., "Efficiency of VHF/FM Antenna-Diversity Reception in Motor Vehicles", Rundfunktechn. Mitteilungen, 31 (1987), p. 221 - 227.
- /9/ Lindenmeier, H. K. and Reiter, L. M., "Multiple-Antenna-Diversity for FM Carradio Reception", IEEE-Antennas and Propagation, Virginia 1987, Conference Proceedings, p. 456 - 459.
- /10/ Lindenmeier, H. K. and Hopf, J. F., "Investigations for determining the Minimum Signal Level necessary for VHF Broadcast Reception in Motor Vehicles", Rundfunktechn. Mitteilungen, 28 (1984), p. 74 - 81.
- /11/ Lindenmeier, H. K., Hopf, J. F. and Reiter, L. M., "Vehicle-integrated Multiple Antenna System for FM-Scanning-Diversity", Conference of the ITG, Section Antennas, Oct. 1988, Lindau, FRG, Conference Proceedings, p. 78 - 80.

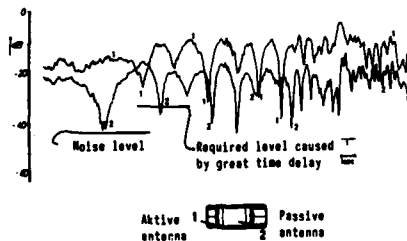


Fig. 1: Measured RF-level in a driven car

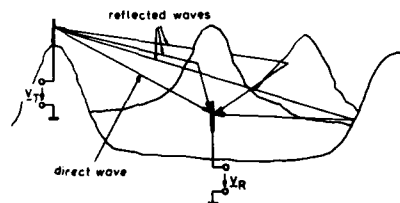
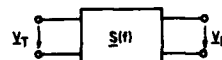


Fig. 2a: Transmission between transmitter and receiver with a direct wave and reflected waves



$$Y_R = Y_T \cdot A_0 \exp(j\omega_T t + j\theta(t)) + \sum_{i=1}^N Y_T \cdot A_i \exp(j\omega_T(t - \tau_i) + j\theta(t - \tau_i))$$

Fig. 2b: Twoport between the transmission- and the receiving antenna

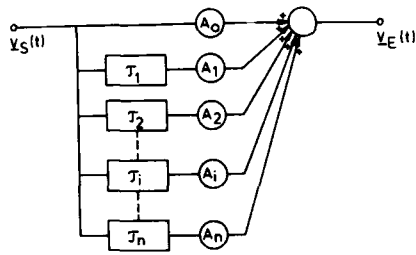


Fig. 3: Structure of a non recursive time delay filter

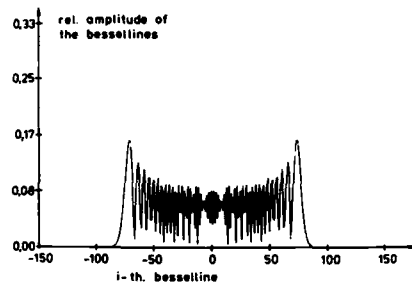


Fig. 5a: Envelope of a Bessel-spectrum of a non disturbed FM-modulated rf-signal

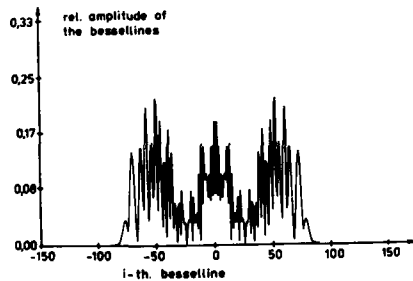


Fig. 5b: Envelope of a Bessel-spectrum of an FM-modulated rf-signal disturbed by multipath reception

Fig. 7: Distortion factor as a function of the AF-frequency and the delay time difference for two superimposed waves, $A_1/A_0=0.5$

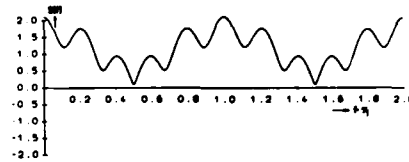


Fig. 4a: Channel transfer function

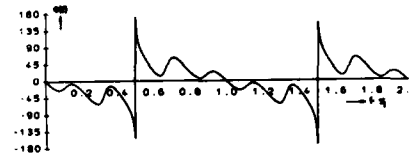


Fig. 4b: Phase function; $A_1/A_0=0.8$, $A_2/A_0=0.4$, $\tau_1=20 \mu s$, $\tau_2=100 \mu s$

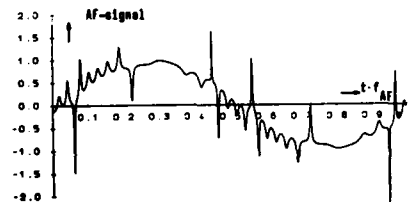


Fig. 6a: Distorted AF-signal at the receiver output

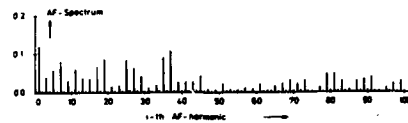
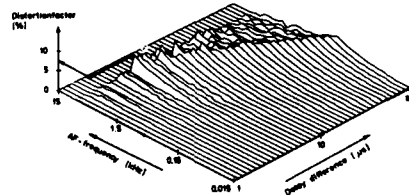


Fig. 6b: Distortion AF-spectrum of disturbed signal in fig. 6a (without desired signal)



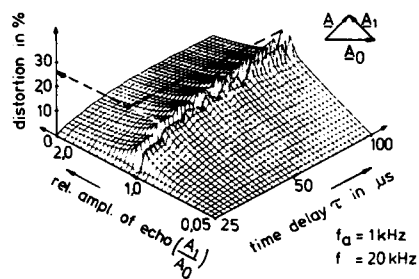


Fig. 7a: Multipath distortions

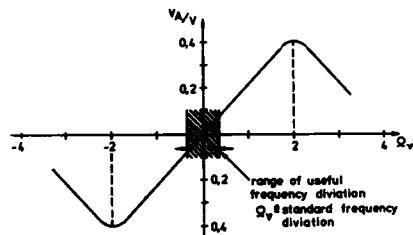


Fig. 10: Demodulation characteristic of the radiodetector

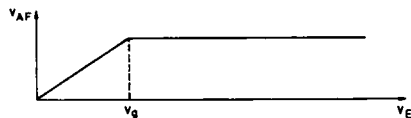


Fig. 11: AF-signal as a function of the RF-level

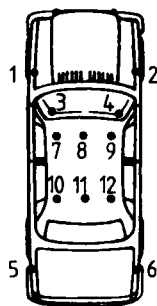


Fig. 13: Car under investigation with 12 antennas.

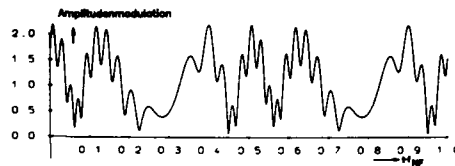
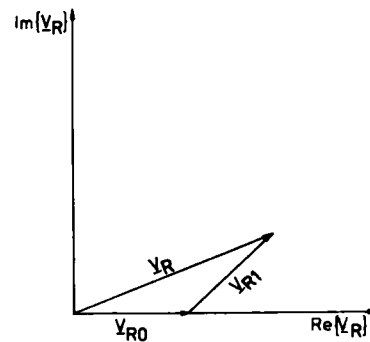
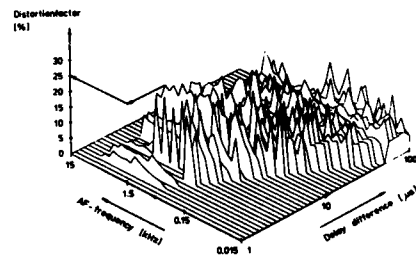
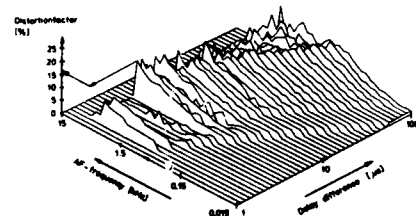
Fig. 8: Related amplitude modulation A/A_0 at the receiver input caused by multipath reception

Fig. 9: Vector diagram for direct and one reflected wave

Fig. 12a: Distortion factor as a function of the AF-frequency and the delay difference for two superimposed waves by application of the ideal FM-demodulator $A_1/A_0=0.9$ Fig. 12b: Distortion factor as a function of the AF-frequency and the delay difference for two superimposed waves by application of the ratio detector, $A_1/A_0=0.9$

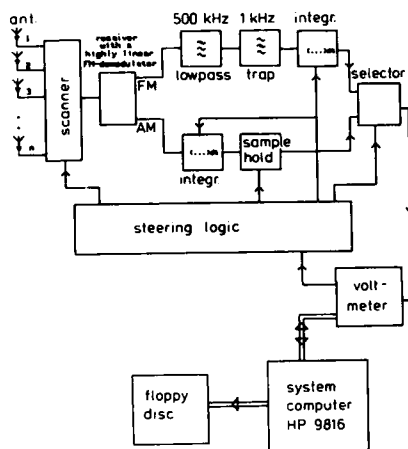


Fig. 14: Measuring device for efficiency evaluation of antenna diversity with various numbers of auxiliary antennas.

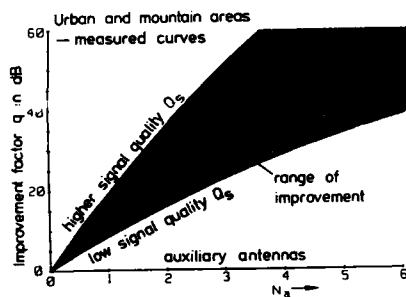


Fig. 19: Region (shaded range) of the expected improvement factor in city areas and mountain areas as well versus the number N_a of auxiliary antennas.

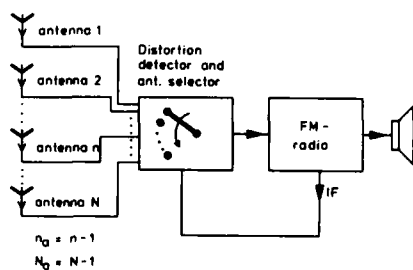


Fig. 20: Principle diagram of realised antenna diversity system with fast indicating distortion detector (detection time $25 \mu s$).

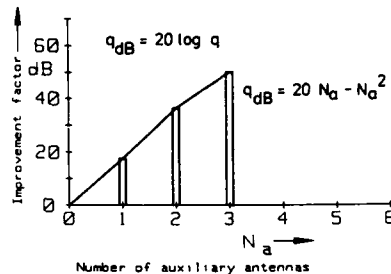


Fig. 15: Logarithmic improvement factor in Munich city area with poor signal supply versus the number N_a of auxiliary antennas.

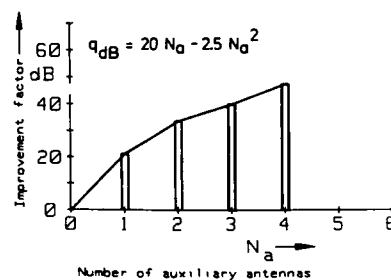


Fig. 16: Improvement factor similar to Fig. 15 however measured at different location in Munich and different day and including a greater N_a .

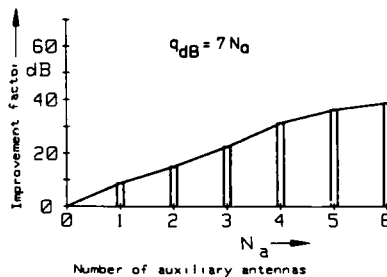


Fig. 17: Improvement factor versus N_a measured in mountain area with great time delay differences and poor signal supply

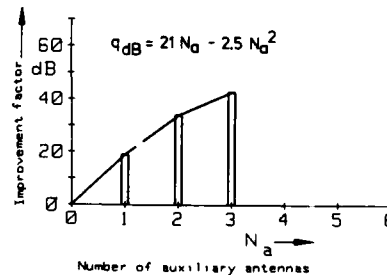


Fig. 18: Improvement factor similar to Fig. 17 however with good signal supply.

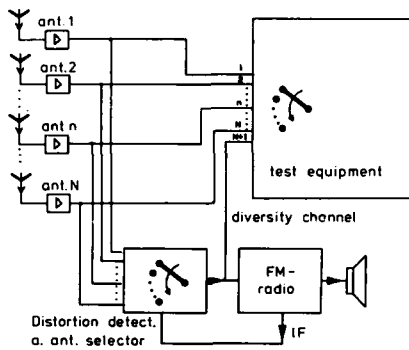


Fig. 21: Test equipment for evaluating the improvement factor of the realised antenna diversity system in comparison with measured and computed results

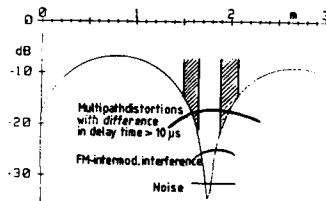


Fig. 27: Different kinds of distortions in a level minimum

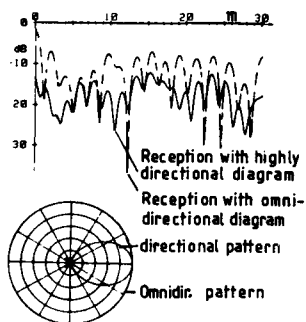


Fig. 28: Variation of signal level with an omnidirectional antenna pattern (---) a. with directional antenna pattern (—)

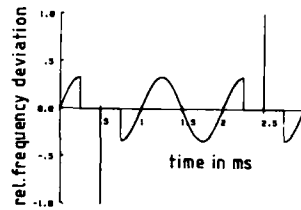


Fig. 22: Testsignal for registration of disturbing echo-waves

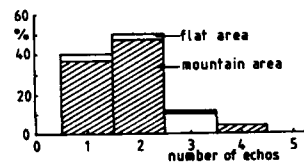


Fig. 23: Probability distribution of the number of echo-waves for an audio distortion greater than 20%

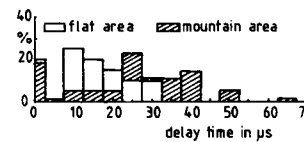


Fig. 24: Probability distribution of the delay time

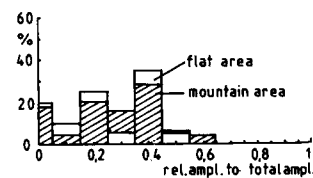


Fig. 25: Probability distribution of the wave amplitudes related to the total amplitude

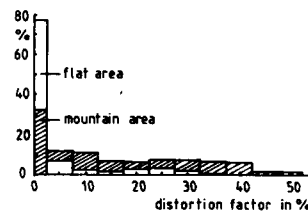


Fig. 26: Probability distribution of the distortion factor

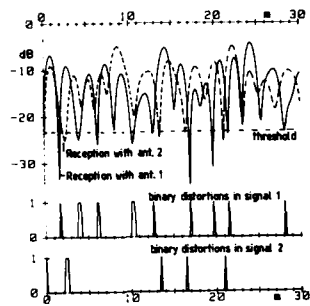


Fig. 29: Variation of signal level of 2 antennas with binary distortion indication by a fixed threshold

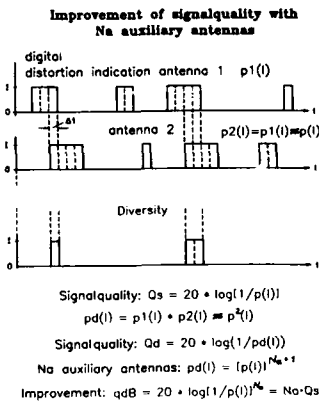


Fig. 29a: Explanation of improvement factor

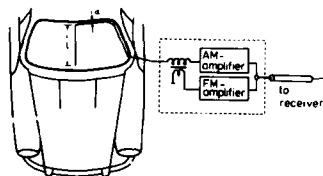


Fig. 33: FM-AM windshield antenna

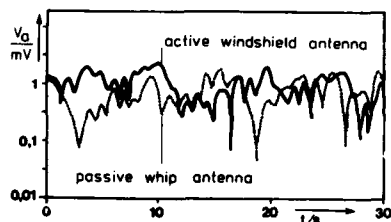


Fig. 35: Rf-level versus time

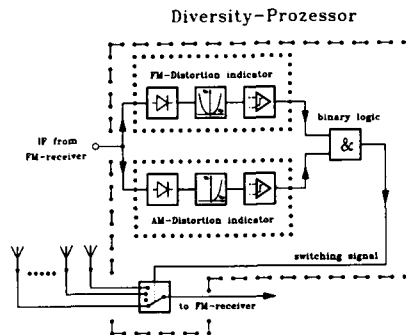


Fig. 30: Diversity processor with AM-FM-indication

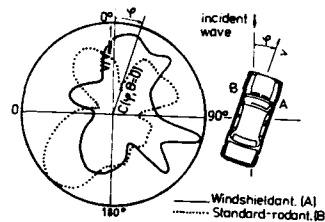


Fig. 31: Horizontal diagram with single wave

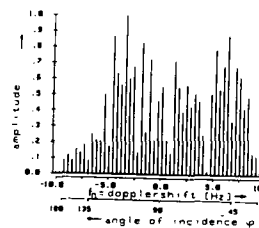


Fig. 32: Amplitudes of waves versus the angle of incidence

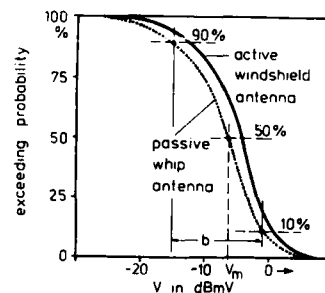


Fig. 36: exceeding probability of rf-level

passive whip antenna in the rear of car in Fig.33

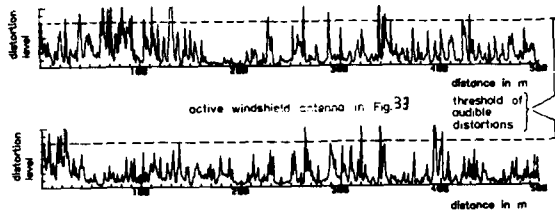


Fig. 34:
Distortion
level versus
driven dis-
tance at the
output of the
fast detecting
distortion in-
dicator.

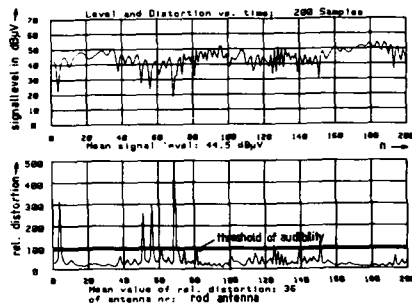


Fig. 37: Signal- and distortion level of
a standard rod antenna

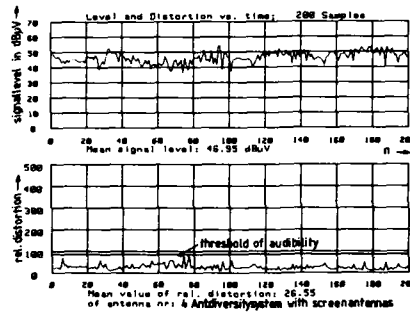


Fig. 38: Signal- and distortion level of
a 4-Ant. diversity system with four screen
antennas (s. Fig. 39).

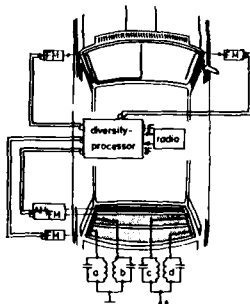


Fig. 39: Antenna diversity system with 4
antennas

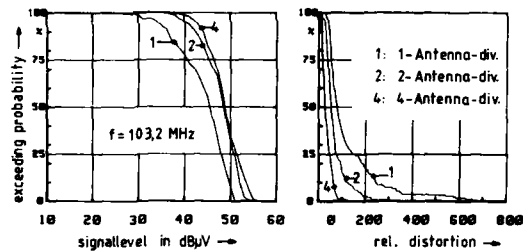


Fig. 40: Exceeding probability of signal
levels and the appertaining distortion
levels with 1, 2 or 4 antennas in the
diversity system

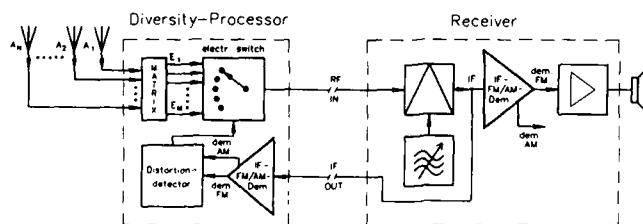


Fig. 41: Diversity system with external diversity processor

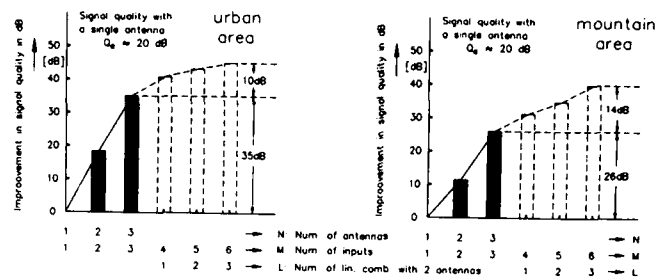


Fig. 42: Improvement factor versus the number of antennas and linear combinations of antenna signals

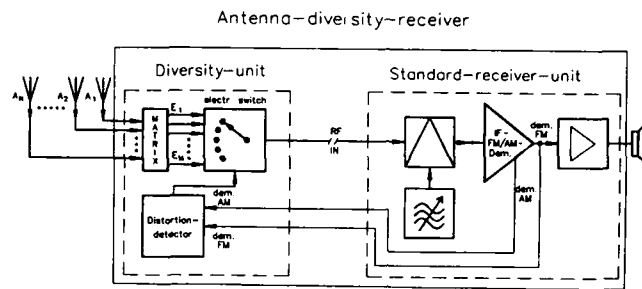


Fig. 43: Diversity system with internal diversity processor

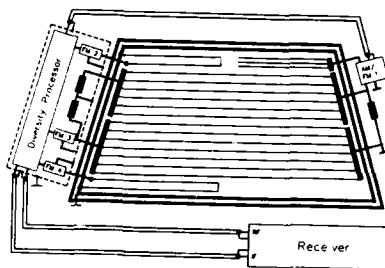


Fig. 44: Integrated antenna system with four FM-antennas and an AM-antenna



Fig. 45: Four mast antennas for FM-diversity

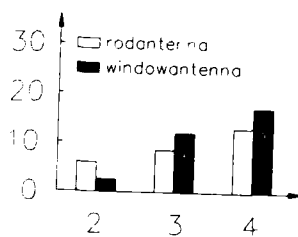


Fig. 46: Improvement factor in dB versus the number of antennas on the rear window

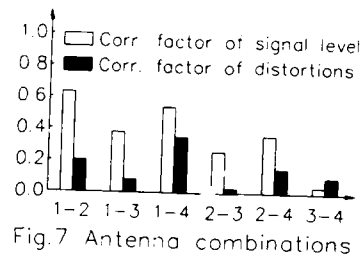


Fig. 47: Correlation factor of the signals of 4 antennas on a window

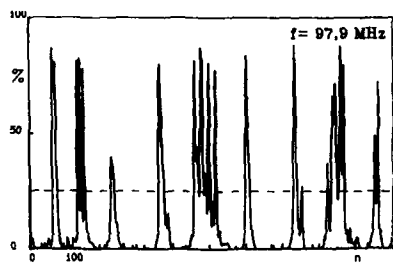


Fig. 48a: Distortion versus time with 1 antenna

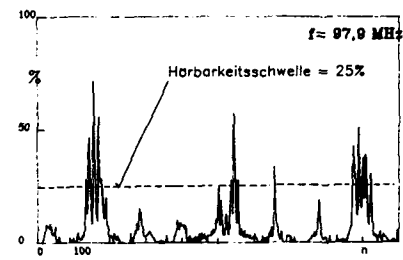


Fig. 48b: Distortion versus time with 2 antennas

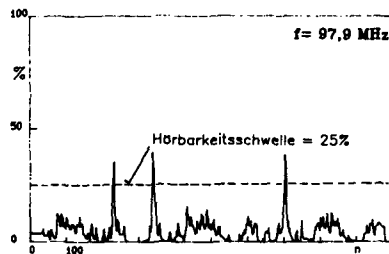


Fig. 48c: Distortion versus time with 3 antennas

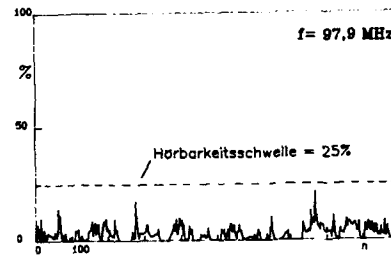


Fig. 48d: Distortion versus time with 4 antennas

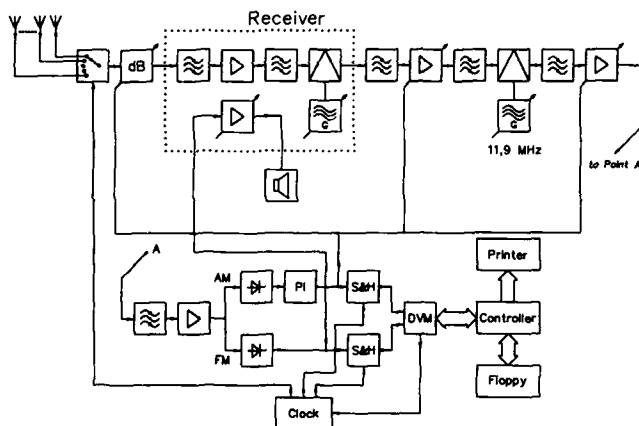


Fig. 49: Measuring equipment with computer for data storage and evaluation

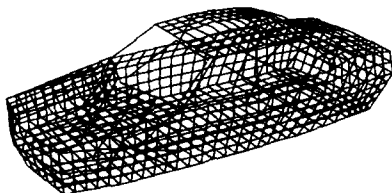


Fig. 50a: Wire grid model of car with windshield antenna for computation

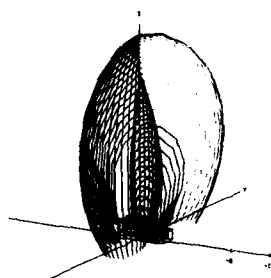


Fig. 50b: Three dimensional radiation diagram of antenna in fig. 50a for horizontal polarization

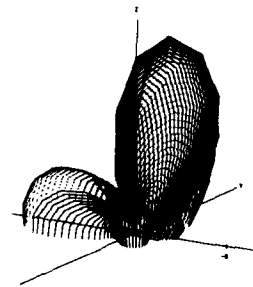


Fig. 50c: Three dimensional radiation diagram of antenna in fig. 50a for vertical polarization

**GTD/UTD : BRIEF HISTORY OF SUCCESSIVE DEVELOPMENT OF THEORY AND RECENT ADVANCES
- APPLICATIONS TO ANTENNAS ON SHIPS AND AIRCRAFT**

by
F.A. Molinet
Société MOTHESIM, La Boursidière
RN 186
92357 Le Plessis-Robinson, France

SUMMARY

The basic concepts of GTD and its relation to the wave equation are reviewed in the frame-work of asymptotic expansions which historically formed the basis for further developments and improvements of the theory. Uniform solutions and especially the Uniform Theory of Diffraction (UTD) are introduced and a brief synthesis of successive development of GTD/UTD is presented. Among the subjects which have been recently advanced, those which either lead to fundamental improvements of the theory or which have an important impact on its applications to current radiation and scattering problems have been selected and analyzed in more details. A great emphasis is given to the extension of GTD/UTD to metallic bodies with coatings. Finally some general ideas for conceiving a computer program founded on GTD/UTD are presented together with some software packages which have been recently developed in France. A few selected applications of these computer programs to predict the EM radiation from antennas on an aircraft or a spacecraft and on a ship are described and illustrated by graphical displays.

1. INTRODUCTION

The Geometrical Theory of Diffraction (GTD) is a ray-optic method valid when the wavelength is small compared to the significant dimensions of the objects with which the field interacts. In spite of the rapid development of modern computers and numerical methods, GTD remains essential for the resolution of scattering problems involving large objects of arbitrary shape like airplanes, helicopters, tanks and ships, at radar frequencies which still exceed the capacity of present machines.

Also GTD has the appearance of a method which is simple to understand and apply, it is frequently employed incorrectly. The reason is that GTD cannot be understood without a minimum of knowledge on perturbation methods and asymptotic expansions. For instance, it is well known that rays are solutions of the eikonal equation. However the domain of validity of that equation is generally not so well known. For example people often think that the eikonal equation is not valid in transition regions like shadow boundaries or caustic regions. But a more careful analysis of the problem, by the boundary layer method which belongs to the class of perturbation methods, shows that this equation remains valid in shadow boundaries but not in caustic regions. Hence we can still use the concept of rays in shadow boundaries, only the transport equation, i.e. the energy flow is modified in these regions when compared to Geometrical Optics. Other concepts which are sometimes used in literature are often misunderstood. For example, the difference between a ray-field and a transition region field is not obvious. As we just mentioned, rays can be used in shadow boundary transition regions. However, it is found that the field in these regions varies rapidly in the direction tangent to the wavefront. People say currently that it has not the behaviour of a local plane wave which is another concept to be clarified. It does not mean only that the form of the wavefront is locally smooth but also that the field along it varies slowly, or equivalently that the transport equation of G.O. holds. A ray field must therefore be understood as a field verifying the laws of G.O. it behaves locally as a plane wave, whereas a transition region field does not verify these laws and may be thought of as a superposition of a continuous spectrum of local plane waves.

Another question which arises in applying GTD to a complex object is the proper choice of the ray species which have to be retained in order to get a good approximation of the diffracted field. Again a correct answer to this question can only be given if we think about the H.F. solution in terms of asymptotic series which are ordered with respect to an asymptotic sequence compatible with the solution we want to represent.

For all these reasons, we have introduced in chapter 2 some elementary definitions and properties of asymptotic series and adopted a point of view in our presentation of GTD which takes into account both aspects of this theory, namely its relation to ray-optics through a generalization of Fermat principle and its relation to wave optics through the concept of asymptotic expansions.

Recent advances in GTD/UTD are presented in chapter 3. Among the selected topics of fundamental importance are the extension of ray theory by including complex specularly reflected incident and creeping rays on concave-convex surfaces, resolution of new canonical problems for surface-edge diffraction and derivation of universal functions for the representation of the field diffracted by a curved wedge in overlapping transition regions. Two other topics which are important in current radiation and scattering problems are also presented : the extension of GTD/UTD to metallic bodies with coatings comprizing some non-published results obtained at MOTHESIM on surface diffraction, and the conception of a general computer program with some French realizations.

A few applications of these programs are presented. They concern the prediction of radiation pattern deformations of antennas mounted on an aircraft and a spacecraft and the evaluation of the coupling between reflector antennas on a ship.

In all the formulas presented in this article, the convention $e^{-i\omega t}$ has been adopted for the time dependance of the electromagnetic field.

2. FUNDAMENTAL CONCEPTS AND BRIEF HISTORY OF SUCCESSIVE DEVELOPMENT OF GTD/UTD

GTD is founded on physical and mathematical concepts having their origin in the work done by Sommerfeld and Runge [1] at the beginning of the century in order to elucidate the relation between ray optics and wave optics. These authors established the first link between Maxwell's equations and ray optics by showing how the laws of optics could be derived from the reduced wave equation. It is important to come back briefly to this important work which enabled later other investigators to extend ray optics to vector fields and to diffraction problems.

In an inhomogeneous medium with refractive index $n(\vec{r})$ where \vec{r} is the position vector, the reduced wave equation can be written as follows :

$$(\Delta + k_0^2 n(\vec{r})) U(\vec{r}) = 0 \quad (1)$$

where $U(\vec{r})$ is the scalar wave function and k_0 the wave number in vacuum, related to the wavelength λ_0 by $k_0 = \frac{2\pi}{\lambda_0}$.

A decisive step in the work of Sommerfeld and Runge was to put the wave equation (1) in a form suitable for the application of the method of perturbations. This was accomplished with the transformation :

$$U(\vec{r}) = e^{ik_0 S(\vec{r})} v(\vec{r}) \quad (2)$$

which transforms the wave equation in the following equivalent equation :

$$(n - |\vec{\nabla} S|^2) + \frac{1}{k_0} (\Delta S + 2 \vec{\nabla} S \cdot \frac{\vec{\nabla} v}{v}) + \frac{1}{k_0^2} \frac{\Delta v}{v} = 0 \quad (3)$$

which has been written in decreasing order of magnitude with respect to the parameter $1/k_0$ which is supposed small.

Now, if we neglect the last term, equation (3) is satisfied when the first two terms are zero. This leads to the well known eikonal and transport equations :

$$|\vec{\nabla} S|^2 = n \quad (4)$$

$$v \Delta S + 2 \vec{\nabla} S \cdot \vec{\nabla} v = 0 \quad (5)$$

It can be shown that the solutions of these equations contain all the laws of ray-optics. We shall not expand on these equations since their solutions are now given in classical textbooks but rather focalize our attention on the third term of (3) which has been neglected. We see that the wave optics analysis given above allows us to establish the condition for the validity of the geometrical-optics approximation, a necessary and sufficient condition being that $k_0^2 \Delta v/v$ is truly negligible. It fails obviously in regions where the derivatives of the amplitude v of the field has a rapid spatial variation within a wavelength which happens for example near a focus or a caustic surface or a shadow boundary. Outside these regions, the third term of (3) is small and can be treated as a perturbation. Since it involves only the amplitude function v and not the phase function S , the perturbation modifies only the function v which can be written formally as a series with respect to powers of the small parameter $\frac{1}{k_0}$:

$$v(\vec{r}) = \sum_{n=0}^{\infty} \frac{1}{k_0^n} v_n(\vec{r}) \quad (6)$$

If we insert (6) in (3) and equal to zero the terms of equal powers of $1/k_0$, we see that the eikonal equation remains the same as (4) and is independent of the order n . However, only the first order term v_0 verifies the transport equation (5), all other terms ($n > 0$) are solution of a system of coupled modified transport equations :

$$v_n \Delta S + 2 \vec{\nabla} S \cdot \vec{\nabla} v_n = - \Delta v_{n-1} \quad (7)$$

which can be solved by iteration if the unperturbed solution v_0 is known. This procedure shows that the ray-optic solution is the first term of a perturbation expansion in powers of $\frac{1}{k_0}$, which is also called an asymptotic series.

The basic characteristic of an asymptotic series which is a consequence of the method of perturbations is the fact that the sum of its first few terms is an approximate solution which becomes more and more exact as the perturbation parameter (here $1/k_0$) becomes smaller and smaller. At the limit when the perturbation parameter tends to zero, the exact solution is given by the first term of the asymptotic series.

The method of perturbations and the notion of asymptotic series have first been introduced by Poincaré in 1892 [2]. Owing to the importance of this subject in the extensions of rayoptics which appeared during the first half of this century and to the prominent part it takes in modern diffraction theory, we give here some elementary definitions and properties of asymptotic series which are essential for a sound understanding of the foundations and successive development of GTD.

First, in the method of perturbations, one does not need to use a power series to represent a function. Instead, one can use a general sequence of functions $v_n(\epsilon)$ of the perturbation parameter ϵ , as long as :

$$v_n(\epsilon) = o(v_{n-1}(\epsilon)) \quad \text{as } \epsilon \rightarrow 0, v_n(\epsilon) > 0 \quad (8)$$

where o is the Landau symbol "small oh". Such a sequence is called an asymptotic sequence. In terms of asymptotic sequences, we can define asymptotic expansions:

$$F(\epsilon) = \sum_{n=0}^N a_n v_n(\epsilon) + o(v_N(\epsilon)) \quad (9)$$

where a_n is independent of ϵ . However, in order that such an asymptotic expansion be an approximate representation of a function $f(\epsilon)$ as ϵ tends to zero, the coefficients a_n must verify:

$$\begin{aligned} a_0 &= \lim_{\epsilon \rightarrow 0} \frac{f(\epsilon)}{v_0(\epsilon)} \\ a_1 &= \lim_{\epsilon \rightarrow 0} \frac{f(\epsilon) - a_0 v_0(\epsilon)}{v_1(\epsilon)} \\ a_K &= \lim_{\epsilon \rightarrow 0} \frac{f(\epsilon) - \sum_{n=0}^{K-1} a_n v_n(\epsilon)}{v_K(\epsilon)} \end{aligned} \quad (10)$$

This has only a meaning if the limits exist which implies some restrictions in the choice of an asymptotic sequence. It is easy to see that if b_n are the coefficients of (9) which are different from zero, the sequence $v_n(\epsilon)$ will be compatible with the function $f(\epsilon)$ if it contains the sequence $\mu_n(\epsilon)$ defined by:

$$\begin{aligned} \text{Ord}[\mu_0(\epsilon)] &= \text{Ord}[f(\epsilon)] \\ \text{Ord}[\mu_1(\epsilon)] &= \text{Ord}[f(\epsilon) - b_0 \mu_0(\epsilon)] \\ &\text{-----} \\ \text{Ord}[\mu_K(\epsilon)] &= \text{Ord}[f(\epsilon) - \sum_{n=0}^{K-1} b_n \mu_n(\epsilon)] \end{aligned} \quad (11)$$

In practice, we do not know the function $f(\epsilon)$ when we apply the method of perturbations since it is the unknown function of the problem. Hence it is not always easy to choose an asymptotic sequence which is compatible with the solution we want to represent. We will see later that this difficulty can be overcome by injecting some a priori knowledge given by known solutions of the wave equation in special configurations.

The asymptotic expansion of a function $f(\epsilon)$ with respect to an asymptotic sequence compatible with $f(\epsilon)$ is unique. However, different compatible sequences define different asymptotic expansions. In consequence of its definition (9), an asymptotic expansion is limited to $N+1$ terms. It is an approximate representation of the function $f(\epsilon)$ as the parameter ϵ tends to zero, the error being $o(v_N(\epsilon))$ since according to (9) (10) and (11):

$$|f(\epsilon) - F(\epsilon)| = o(v_N(\epsilon)) \quad (12)$$

If it is possible to construct by recurrence all the terms of an asymptotic expansion when N tends to infinity, then we get an infinite series which we call an "asymptotic series". An asymptotic series $F(\epsilon)$ is therefore an expansion of a given function $f(\epsilon)$ in an infinite series with respect to a compatible asymptotic sequence $v_n(\epsilon)$. It is a purely formal expansion which we write:

$$F(\epsilon) = \sum_{n=0}^{\infty} a_n v_n(\epsilon) \quad (13)$$

This series may be convergent or divergent. When it is divergent we can still give a meaning to (13) by truncating the series after $N+1$ terms and write:

$$|f(\epsilon) - F_N(\epsilon)| = o(v_N(\epsilon)) \quad \text{as } \epsilon \rightarrow 0 \quad (14)$$

where $F_N(\epsilon)$ is an asymptotic expansion defined by (9).

A direct consequence of (14) is that for a fixed value N_1 of N and a given positive number η which can be chosen arbitrarily small, we can always find a value ϵ_1 of ϵ such that:

$$\frac{|f(\epsilon_1) - F_{N_1}(\epsilon_1)|}{v_{N_1}(\epsilon_1)} < \eta \quad (15)$$

Conversely, if we fix $\epsilon = \epsilon_1$ and augment N from 0 to infinity, the difference $|f(\epsilon_1) - F_N(\epsilon_1)|$ will pass through a value smaller than $\eta v_{N_1}(\epsilon_1)$ for $N = N_1$ and then augment indefinitely with N if the series (13) is divergent. Hence, for a fixed value of ϵ there is an optimum value of N for which the difference $|f(\epsilon) - F_N(\epsilon)|$ is minimum. Unfortunately, no general rule exists which permits to predict the position of this optimum and some experience is needed in fixing the number of terms in an asymptotic expansion. However this is not a too severe problem since in most practical situations it is not possible to construct more than the first two terms of an asymptotic expansion. The latter is therefore useful for ϵ sufficiently small for which it is generally much more efficient for numerical

evaluations of the solution than expansions that are uniformly and absolutely convergent.

With these concepts in mind, we can now follow the historical development of ray optics and GTD.

The work of Sommerfeld and Runge was later extended to vector fields [3, 4] and has given rise to the modern Geometrical Optics (G.O.). For vector fields in an homogeneous and isotropic medium, the scalar wave equation is replaced by the vector Helmholtz equation :

$$\Delta \vec{E} + k^2 \vec{E} = 0 \quad (16)$$

where \vec{E} satisfies furthermore the Gauss law $\nabla \cdot \vec{E} = 0$.

Hence, by using the same transformation as in (2), with the scalar wave function $v(\vec{r})$ replaced by the vector wave function $\vec{E}(\vec{r})$, we get :

$$(n - |\vec{\nabla} S|^2) \vec{E} + \frac{1}{k_0^2} (\Delta S + 2 \vec{\nabla} S \cdot \nabla) \vec{E} + \frac{1}{k_0^2} \Delta \vec{E} = 0 \quad (17)$$

with :

$$\vec{\nabla} S \cdot \vec{E} + \frac{1}{k_0} \vec{\nabla} \cdot \vec{E} = 0 \quad (18)$$

Equation (17) is similar to (3) and can be submitted to the same treatment. We see that if we neglect the last term of (17), we find the same eikonal equation as for a scalar field. The transport equation is also the same as in (5) for each component of the vector field. Equation (18) gives us an information concerning the polarization. At first order in $1/k_0$, we have $\vec{\nabla} S \cdot \vec{E} = 0$. Since $\vec{\nabla} S$ is orthogonal to the wave front $S(\vec{r}) = Ct$ we see that \vec{E} is tangent to the wave front and orthogonal to the ray trajectory. The same is true for the corresponding magnetic field \vec{H} which, in addition, by using the second Maxwell equation, is found to be orthogonal to \vec{E} at first order in $1/k_0$.

These properties associated with the solutions of the eikonal and transport equations generalize the laws of optics to vector fields and constitute the modern Geometrical Optics. They have also been extended to an inhomogeneous isotropic medium by observing that when the wavelength is so small that significant medium changes occur only over distances large compared with it, a reasonable assumption is that in local regions the field behaves as if it were in an homogeneous medium.

Furthermore, by applying the method of perturbations to (17), or its equivalent for an inhomogeneous medium, it can be shown that the G.O. field is the first term of an asymptotic series with respect to the asymptotic sequence k_0^{-n} . This series is also called the Luneburg-Kline series and is usually written in powers of $(ik_0)^{-n}$:

$$\begin{aligned} \vec{E}(\vec{r}) &\approx e^{ikS(\vec{r})} \sum_{n=0}^{\infty} \frac{1}{(ik_0)^n} \vec{E}_n(\vec{r}) \\ \vec{H}(\vec{r}) &\approx e^{ikS(\vec{r})} \sum_{n=0}^{\infty} \frac{1}{(ik_0)^n} \vec{H}_n(\vec{r}) \end{aligned} \quad (19)$$

A second significant advance in ray optics was its extension to diffraction problems. This development, due mainly to J.B. Keller [5, 6, 7] in the early 1950 greatly increased the domain of application of ray methods and has become known as the Geometrical Theory of Diffraction (GTD).

Keller's theory is usually presented as an heuristic extension of Geometrical Optics and stated in the form of postulates. Although this way to introduce GTD highlights some useful aspects of GTD, it obscures its relation to the wave equation which forms the basis for further developments and improvements of the theory. We adopt therefore a point of view which takes into account both aspects through the concept of asymptotic expansions already introduced to justify G.O.

Scattering occurs when an obstacle is placed in the field. When this obstacle is large in terms of a wavelength, scattering is found to be essentially a localized phenomenon identifiable with certain localized areas on the object : points of specular reflection, shadow boundaries, edges and vertices. At high frequencies, the scattered field may therefore be thought of as being the sum of distinct contributions associated with different rays comprising the reflected rays and various species of diffracted rays. Diffracted rays are produced by incident rays which hit edges, vertices (corner or tip) of the boundary surface or which graze such a surface. In order to determine these rays Keller proposed a generalization of Fermat's principle to include any path satisfying constraints on the surface of the obstacle and having stationary optical length. For example, an edge diffracted ray from P to Q is a curve which has stationary optical length among all curves from P to Q having a point on an edge. In an homogeneous medium, this implies that the ray is a straight line from P to a point on the edge and a straight line from that point to Q. The two lines make equal angles with the edge at the point where they meet it and lie on opposite sides of the plane normal to the edge at the point of diffraction. This is the law of edge diffraction. It shows that one incident ray produces a cone of diffracted rays (Fig. 1). The laws of edge diffraction apply to higher order edge diffracted rays produced by higher order edges, which are lines of discontinuities of the curvature or some higher order derivative on the boundary surface.

Another important class of diffracted rays are creeping rays, which appear when incident rays strike a smooth, curved boundary surface at grazing incidence, i.e. at the shadow boundary. In an homogeneous medium, a ray incident on the shadow boundary at Q, divides into two parts: one part of the incident energy continues straight on as predicted by geometrical optics and a second part follows the surface into the shadow region as a creeping ray, which

sheds diffracted rays tangentially as it propagates. Surface diffracted rays follow the modified Fermat principle for surface diffraction: a surface-diffracted ray from a point P to a point Q is a curve which makes stationary the optical length among all curves from P to Q having an arc on the boundary surface. According to this law, the incident ray and the resulting surface diffracted ray in the same medium, are parallel to each other at the point of diffraction and lie on opposite sides of the plane normal to the ray at this point. When the two rays lie in different media, they obey the law of refraction. In an homogeneous medium, the arc Q_1Q_2 followed by the ray on the boundary is an arc of a geodesic or shortest path on the surface (Fig. 2) and the incident and diffracted rays are tangent to the surface ray at Q_1 and Q_2 .

By combining the law of edge diffraction with the law of surface diffraction, a new class of rays called edge excited creeping rays can be defined. When an incident ray hits an edge in a curved surface, two creeping rays are launched in the directions of the generatrices of the Keller cone tangent to the surfaces adjacent to the edge (Fig. 3).

Finally, an incident ray which hits a vertex (corner or tip) in an homogeneous medium produces straight diffracted rays which leave the vertex in all directions (Fig. 4). This is the law of vertex diffraction. Higher order diffracted rays produced by higher order vertices, which are discontinuities in higher order derivatives of an edge or a higher order edge follow also the law of vertex diffraction. Keller justified his generalization of the concept of rays through the geometrical interpretation of the asymptotic expansion of exact solutions of the wave equation. For some special boundaries, the wave equation can be solved exactly and the first few terms of the asymptotic expansion of the solution can be derived. Among them are the perfectly conducting wedge and the perfectly conducting cylinder or sphere, illuminated by a plane wave. Keller showed that the asymptotic solutions of the wave equation for these special boundaries have a simple geometrical interpretation in terms of edge diffracted rays and creeping rays which verify respectively the laws of edge and surface diffraction.

Instead of extending the concept of rays, one can also extend the asymptotic sequence in order to get a sequence which is compatible with the solution of the diffraction problem we want to represent. From the asymptotic expansions of the exact solutions of the half-plane and the wedge, we see that the asymptotic sequence k_0^{-n} is not compatible with these solutions. In order to get a compatible asymptotic sequence, we must add the sequence k_0^{-n-1} . Moreover, the complete asymptotic expansion is found to be the sum of two separate contributions, each of which is represented by a different asymptotic series: an asymptotic series defined with respect to the asymptotic sequence k_0^{-n} corresponding to the reflected field and an asymptotic series defined with respect to the asymptotic sequence k_0^{-n-1} corresponding to the edge diffracted field. The complete asymptotic series representing the solution of a general diffraction problem may therefore be thought of as being the sum of different subseries corresponding to different diffraction phenomena.

From asymptotic expansions of exact solutions of the wave equation, the following correspondence between diffraction phenomena and compatible asymptotic sequences have been established:

<u>Diffraction phenomena</u>	<u>Asymptotic sequence</u> $v_n(\frac{1}{k_0})$, $n = 0, 1, 2, \dots$
reflection by a smooth surface	k_0^{-n}
diffraction by a sharp edge	$k_0^{-n-1/2}$
diffraction by a curvature discontinuity	$k_0^{-n-3/4}$
diffraction by a vertex	k_0^{-n-1}

We see that the asymptotic series defined by these asymptotic sequences can all be written in the form of a Luneburg-Kline series (19), multiplied by the factor $v_0(\frac{1}{k_0})$ which by definition is equal to the first term of the corresponding asymptotic sequence. For example $v_0(1/k_0) = 1$ for the field reflected by a smooth surface and $v_0(1/k_0) = k_0^{-1/2}$ for the field diffracted by a sharp edge. Each of the diffraction phenomena of the preceding list can therefore be written:

$$\vec{A}(\vec{r}) = v_0(\frac{1}{k_0}) e^{ikS(\vec{r})} \sum_{n=0}^{\infty} \frac{1}{(ik_0)^n} \vec{a}_n(\vec{r}) \quad (20)$$

where \vec{A} represents either the electric field or the magnetic field.

It can be shown that solutions of the form (20) are not the most general forms of asymptotic expansions encountered in diffraction problems. The exact solutions of the diffraction of a plane wave by a circular cylinder and a sphere show that the asymptotic series representing the field diffracted in the deep shadow of a smooth convex surface has the form:

$$\vec{A}(\vec{r}) = e^{ikS(\vec{r})} - k_0^\alpha \chi(\vec{r}) \sum_{n=0}^{\infty} (ik_0)^{\lambda_n} \vec{a}_n(\vec{r}) \quad (21)$$

where $\lambda_n = \frac{n}{3}$ and $e^{-k_0^\alpha \chi(\vec{r})}$ is a decay factor with $\alpha = 1/3$. General solutions of the form (21) have first been constructed by Friedlander and Keller [8] for the scalar wave equation. The series (21) is therefore also called the Friedlander-Keller series.

In order that the asymptotic series (20) and (21) be solutions of Maxwell equations, they must verify the eikonal and transport equation obtained by substituting these series into the vector Helmholtz equation (16). In addition, they must verify the Gauss law.

Since (20) has the same form as (19) the equations verified by the phase function $S(\vec{r})$ and the vectorial amplitudes \vec{a}_n and \vec{h}_n are the same, whatever the diffraction phenomenon considered may be, and are identical to those verified by the incident field which for an

homogeneous medium with refractive index N and characteristic impedance $\eta = \sqrt{\mu/\epsilon}$ are found to be :

$$\begin{aligned} |\vec{\nabla} S|^2 &= N \\ (\Delta S + 2 \vec{\nabla} S \cdot \vec{\nabla}) \vec{e}_n &= -\Delta \vec{e}_{n-1} \\ \vec{\nabla} S \cdot \vec{e}_n &= -\vec{\nabla} \cdot \vec{e}_{n-1} \\ \vec{h}_n &= \frac{1}{\eta} \vec{\nabla} S \times \vec{e}_n + \vec{\nabla} \times \vec{e}_{n-1} \\ n &= 0, 1, 2, \dots; \vec{e}_{-1} = 0 \end{aligned} \quad (22)$$

where \vec{h}_n has been related to \vec{e}_n by the second Maxwell equation. The phase function $S(\vec{r})$ verifies therefore the eikonal equation of G.O. Furthermore, the leading term of each asymptotic subseries verifies the transport equation of G.O. and according to Gauss law, its polarization has the same behavior as that of the G.O. field. Consequently, the diffracted field associated with each diffraction phenomenon represented by (20) propagates along rays which are orthogonal to the wave fronts $S(\vec{r}) = \text{constant}$ and the leading term of the corresponding asymptotic series verifies the normal laws of G.O. In order to determine completely the solution of (22), we must relate the scattered field to the incident field by applying the boundary conditions on the surface of the obstacle. We consider first the eikonal equation which is independent of the transport equations. It is a first order partial differential equation the solution of which is completely determined if we impose some complementary conditions. For example, we can impose the values of the unknown function on a regular surface which may be an initial wave front or the regular part of the illuminated surface of the body. This procedure is adequate for the reflected field since the values of the phase function $S^R(\vec{r})$ of the family of reflected rays, at regular points of the illuminated surface of the body are related to those of the phase function $S^I(\vec{r})$ of the incident ray family by the boundary conditions and are therefore known. But we can also give one of the branches of the envelope or caustic of the unknown ray trajectories since the knowledge of this envelope allows us to reconstruct the whole ray family and consequently the phase functions $S(\vec{r})$. When one branch of the caustic is degenerated into a curved line or a focal point, mixed conditions which consist in specifying the equation of the degenerated caustic and in giving the values of the unknown function on it, may be imposed. This last procedure is adequate for edge or vertex diffraction since the edge and the vertex are degenerated caustics for edge and vertex diffracted rays respectively. Finally we see that $S(\vec{r})$ is completely determined for each subseries (20) if we impose a matching condition for the phase at the points where each ray species interacts with the surface of the body. The simplest matching condition is the continuity of the phase :

$$S^I(\vec{r}) = S^R(\vec{r}), \quad \vec{r} \in S_e \quad (23)$$

$$S^I(\vec{r}) = S^d(\vec{r}), \quad \vec{r} \in c \text{ or } \vec{r} = Q \quad (24)$$

where S_e is the regular part of the illuminated region of the body, and where c and Q are respectively an edge and a focal point. Condition (23) is a consequence of the boundary conditions whereas condition (24) is justified by the asymptotic expansions of known asymptotic solutions. It can be shown that both conditions are equivalent to the Fermat principle extended to reflection, edge diffraction and vertex diffraction and contain the corresponding laws.

We now consider the transport equations in (22). It can be shown that these equations reduce to a system of coupled linear differential equations along a ray. In order to determine completely the solution, we must impose the value of \vec{e}_n for $n = 0, 1, 2, \dots$ at one point on each ray. If we choose the point where the ray intersects the surface of the body we can relate \vec{e}_n to the corresponding component of the incident field \vec{e}_n^i at that point by applying the boundary conditions. However this is only possible for the reflected rays since for edge or vertex diffracted rays, the point of interaction lies on a caustic where the field predicted by the general solution of the transport equations in ray coordinates is infinite. It is possible to avoid this problem by observing that the amplitude of the field near an edge behaves like $1/\sigma$ where σ denotes the distance along a ray measured from its point of intersection with the edge. Hence, since the limits :

$$\vec{\delta}_n = \lim_{\sigma \rightarrow 0} (\vec{e}_n(\sigma) \sqrt{\sigma}) \quad (25)$$

exist for all n , we can relate $\vec{\delta}$ instead of \vec{e}_n to the corresponding component $\vec{e}_n^i(0)$ of the incident field on the edge.

Furthermore since we are dealing with linear phenomena, $\vec{\delta}_0$ must be proportional to the leading term of the incident field on the edge. The constant of proportionality is referred to as a diffraction coefficient and for an electromagnetic field it is a dyadic. Hence we can write :

$$\vec{\delta}_0 = \vec{e}_0^i(0) \cdot \vec{D}_e \quad (26)$$

where \vec{D}_e is the dyadic edge diffraction coefficient.

A similar result holds for a vertex where the field behaves like $1/\sigma$ and where we have to consider the limits :

$$\vec{q}_n = \lim_{\sigma \rightarrow 0} (\vec{e}_n(\sigma) \sigma) \quad (27)$$

which for $n = 0$ are proportional to $\vec{e}_0^i(0)$, the constant of proportionality being the dyadic

vertex diffraction coefficient.

The diffraction coefficients depend on the properties of the diffracted field close to the edge or the vertex. In these regions, owing to the presence of the caustic, the amplitude of the diffracted field varies rapidly and the assumptions made in deriving the asymptotic series (20) by the method of perturbations are no longer valid. We need therefore another solution valid close to the edge or the vertex. Two methods may be followed and have been applied in the past. The first one calls upon the local nature of the high frequency diffraction and makes use of the tangent plane approximation. The second method consists in applying the boundary layer theory. This method is more general than the first one and gives a mathematical justification of it. It has first been introduced in diffraction problems by Buchal and Keller [9]. A brief description of the method can be found in [10]. The tangent plane approximation is best defined in the case of a sharp edge in a curved surface. It consists in replacing the surfaces on both sides of the edge, at the point of diffraction by tangent planes forming an infinite straight wedge. In addition, in the vicinity of the point of diffraction, the incident wave is replaced by a plane wave. Thus, the initial problem is reduced to a simpler problem also called canonical problem, namely the diffraction of a plane wave by a straight wedge, the exact solution of which is known. By expanding this solution asymptotically and comparing the results with that of (26), the dyadic diffraction coefficient can be determined. This procedure is the standard method of GTD developed by Keller. It gives only δ_0 . The values of δ_n for $n > 0$ depend not only on the amplitudes of the incident and diffracted waves, but also on their first and higher-order derivatives at the point of diffraction, which depend themselves on higher-order derivatives of the surface of the scatterer in the vicinity of that point. They can therefore not be obtained by the tangent plane approximation.

It can be shown, in a similar way, that expansion (21) leads to a theoretical foundation of GTD for creeping waves. Indeed, by substituting (21) into (16) it is found that the phase function $S(\vec{r})$ verifies the eikonal equation of GO. The solution of this equation is completely defined by specifying the surface of the body which is a caustic for the creeping rays and by imposing the continuity of the phase on the shadow boundary curve where the incident rays are tangent to the body. Furthermore, in an homogeneous medium it is found that the equal-amplitude surfaces $\chi(\vec{r}) = \text{constant}$ are orthogonal to the equal-phase surfaces $S(\vec{r}) = \text{constant}$ and are therefore generated by the rays which are straight lines orthogonal to the same surfaces. Hence, the value of $\chi(\vec{r})$ is constant on each ray. It can also be shown that the amplitude vectors \vec{a}_n verify linear ordinary differential equations along the rays, similar to (22) and that \vec{a}_0 verifies the transport equation of GO. However, the right hand-side of the equations verified by the higher order terms are different and depend on the derivatives of the function χ . Finally, the arbitrary elements in the construction of the solution are the initial values of \vec{a}_n and the value of χ on each ray. These quantities may be adjusted at least for $n = 0$ in order that the expansion corresponds to the solution of a particular boundary value problem.

Explicit expressions for the leading term of the field diffracted by the edge of a curved wedge or by a smooth convex surface may be found in the original work of Keller and his co-workers [5, 11] or in the work of Kouyoumjian and Pathak [12, 13, 14]. A short analysis of all available diffraction coefficients is given in [10]. For details on canonical problems see [33]. In the presentation of GTD given above, the boundary conditions do not enter explicitly in the general results obtained which remain therefore valid for general boundary conditions, applied to the interface between an obstacle and the space surrounding it. Only the diffraction coefficients have to be adapted by solving canonical problems taking into account the actual boundary conditions. However, new ray species may occur like refracted rays penetrating into transparent bodies or surface rays on coated metallic bodies, which have to be added to those already considered here.

The asymptotic expansions introduced so far are not uniform. They break down near caustics and in the transition regions adjacent to shadow and reflection boundaries of incoming rays. These shortcomings of GTD have been recognized in the early state of development of the theory. They have been progressively surmounted by the construction of uniform expansions with the aid of boundary layer theory or by direct generalization of uniform asymptotic expansions of exact solutions of canonical problems. This last procedure is generally much simpler since it does not involve stretched boundary layer coordinates, but its complexity grows rapidly if a large number of caustic regions and transition regions is to be covered. In addition, since the asymptotic expansion depends on the asymptotic sequence used, different uniform expansions may be obtained for the same canonical problem. This happened for the wedge problem for which two different asymptotic solutions have been developed: the UTD (Uniform Theory of Diffraction) solution [12] and the UAT (Uniform Asymptotic Theory) solution [15] [16] [17].

It is well known that the exact solution of the diffraction problem of a plane wave by a straight perfectly conducting wedge can be put in the form of an integral along two paths in a complex plane known as the Sommerfeld contours. By deforming these contours in steepest descent paths, it can be shown that the initial problem reduces to the asymptotic evaluation of integrals of a slowly varying function $f(z)$ multiplied by a phase term having a stationary point z_0 near a pole z_1 of $f(z)$. Two methods have been proposed in the past for deriving a uniform asymptotic expansion of such integrals. The method of Pauli-Clemonow [18] [19] consists in writing $f(z)$ in the form of a regular function $g(z)$ multiplied by $(z - z_0)^{-1}$ and in expanding $g(z)$ in a Taylor series in the vicinity of a stationary point. This method which has been followed by Kouyoumjian and Pathak [12] in the derivation of their uniform solution for a perfectly conducting straight wedge, leads to a universal function involving a Fresnel integral, which is discontinuous through the shadow boundaries of the direct and the reflected GO rays and which compensates exactly the discontinuity of the GO field at these boundaries in such a way that the total GO field be continuous there. By enforcing this property in the case of a curved wedge, they extended the argument of their transition function to a perfectly conducting curved wedge and established a uniform solution for that geometry which they called the Uniform Theory of edge Diffraction (UTD). A second method proposed by Oberhettinger and Van der Waerden [20] [21] consists in writing

$f(z)$ in the form of a sum of a regular function $h(z)$ and the singular term $(z-z_0)^{-1}$ and in expanding $h(z)$ in a Taylor series in the vicinity of the stationary point. This method leads to a result which can be expressed as the sum of the non uniform expansion of the integral and a transition function involving the difference between a Fresnel integral and its asymptotic expansion. Both, the transition function and the non uniform expansion are singular when z_0 approaches z_s , but the singularities compensate each other. This procedure leads to UAT. It is found that the argument of the transition function is directly related to the detour of the ray path passing through the edge compared to the direct or reflected ray paths. This geometrical interpretation enabled Ahluwalia [16] and later Lee and Deschamps [17] to generalize the solution to a curved edge in a curved sheet and to a curved wedge, both being perfectly conducting. The two methods mentioned above lead to different uniform expansions for the exact solution of the diffraction of a plane wave by a perfectly conducting straight wedge, since they have been ordered with respect to different asymptotic sequences. However, if each of them is reexpanded with respect to the same asymptotic sequence $k^{-n-\frac{1}{2}}$, they are identical. In the expansion proposed by Oberhettinger-Van der Weerden, the influence of the pole singularity is completely taken into account in the leading term of the expansion whereas the same information is distributed in all the terms of the Pauli-Clemmow expansion. In its initial version, only the leading term of the expansion was retained in UTD. However, when the amplitude of the incident or the reflected GO field vary rapidly at the edge, in the direction perpendicular to the incident or reflected rays, the second term of Pauli-Clemmow's expansion is no longer of negligible value, especially in the transition regions as shown in [17]. Hence a complementary term has to be added to the UTD solution. Hwang and Kouyoumjian [22] established an expression of this term which they wrote in the format of GTD by introducing a new diffraction coefficient called "Slope Diffraction Coefficient". A detailed comparison between the UAT and UTD solutions has been carried out by Boersma and Rhamat-Samii [23] for a specific test problem, namely the diffraction of an arbitrary cylindrical wave by a half-plane which can be solved exactly in terms of a spectral integral representation. One of their final conclusions was that both UAT and UTD, the latter being augmented by the slope diffraction term, provide very similar, but not identical numerical results. When compared with the values of the spectral integral, the differences observed are of no practical importance. The uniform UTD and UAT solutions mentioned so far are only valid in the regions 1, 2 and 3 of fig. (5). In the transition regions of the diffracted rays (regions 4 and 5), and in the deep shadow of those rays, surface diffraction cannot be neglected. In these regions, the field may be described by hybrid diffraction coefficients [24] expressed in terms of products of edge diffraction coefficients and launching coefficients of creeping waves. Expressions for the diffraction coefficients associated with edge-excited creeping waves have also been recently obtained by Idemen and his co-workers [25] who solved rigorously new canonical problems constituted by cylindrical or spherical sheets.

Uniform solutions have also been established in caustic region. Ludwig [26] first succeeded in the construction of a uniform expansion at a smooth convex caustic and a cusped caustic. The universal function appearing in his expansion for a smooth caustic is the Airy function and its derivative. Ludwig's solution reduces to the GO solution away from the caustic and remains finite at the caustic. Numerical results show that in the neighbourhood of a regular caustic, the GO field deviates from the exact solution only in a small angular region close to the caustic. A good approximation may therefore be obtained by linear interpolation between two points on both sides of the caustic where GTD applies. Near a line caustic or a focal point however, where the field is strongly affected by the presence of the caustic, this procedure is not applicable. In these regions other methods have been developed. Basically, these methods consist in finding an integral form of the solution, having the same asymptotic expansion as the solution given by GO or GTD away from the caustic. At observation points located on the caustic and in its neighbourhood, the field is determined by a numerical evaluation of that integral. A physical problem leading to an asymptotic solution having the same form as the field diffracted by a straight edge, is the radiation of a line current. Hence equivalent electric and magnetic line currents can be determined which radiate a field being asymptotically identical to the GTD field away from the caustic and giving the correct behaviour of the field at the caustic. This method is known as the Equivalent Current Method (ECM). Similarly the Physical Optics (PO) approximation which consists in replacing the currents on a surface by GO currents, is useful for calculating the fields at the caustics of the GO rays. It is important to note that ECM which employs GTD to calculate its equivalent sources is a valid procedure for evaluating the field only if the caustic region is not close to the GO shadow boundary transition regions, otherwise we may have to resort to the Physical Theory of Diffraction (PTD). The latter constitutes an extension of PO and was developed originally by Ufimtsev [27]. The basic idea underlying this method consists in improving the GO current near edges by including a correction or fringe current. The field radiated by these currents is obtained by indirect considerations involving an asymptotic high frequency analysis of the PO integral. An updated bibliography on ECM and PTD together with improvements of these theories may be found in [28].

Transition regions of the field diffracted by a perfectly conducting smooth convex surface have also been investigated by many authors in the last twenty years. We can distinguish two boundary layers: a shadow boundary layer 1 near the shadow boundary SB of the direct and reflected rays and a caustic boundary layer 2 of the creeping waves near the surface of the scatterer (Fig. 6). These boundary layers have a common domain in the neighbourhood of the curve Γ separating the illuminated side from the shadow side of the scatterer, where the high frequency solution is extremely complicated. Similar transition regions are found for the field radiated by a source on a smooth convex surface (Fig. 7). A brief literature survey of these subjects is given in [10]. Uniform asymptotic solutions which are very convenient for engineering applications have been derived for a perfectly conducting scatterer with a source on its surface or far removed from it, by Pathak and Kouyoumjian and their co-workers [13,14,22]. In chapter 3 more details will be given on the special Fock

functions appropriate to the transition region and on their extension to coated metallic bodies.

3. RECENT ADVANCES

GTD is still continuing to be improved and completed. Among the subjects which have been recently advanced we consider those which either lead to fundamental improvements of the theory or have an important impact on its applications to actual radiation and diffraction problems. In the first class, we have selected three topics: extension of ray theory by including complex specularly reflected incident and creeping rays in addition to the conventional real rays; resolution of new canonical problems involving edge diffracted creeping rays and edge diffracted whispering gallery modes; construction of new uniform asymptotic expansions in overlapping transition regions. In the second class, we have selected two topics: extension of GTD to bodies with coatings and conception of a general computer program founded on GTD. We will only briefly discuss the first class of topics and extend more on the second class and especially on the radiation of antennas on coated metallic bodies.

Extension of ray theory to reflection from concave-convex surfaces

It has been shown recently by Ikuno and Felsen [29], that a complete asymptotic theory of high frequency scattering from smooth targets requires consideration of real as well as complex ray fields. The authors have shown that when the surface of an obstacle has a concave-to-convex inflection point, the contribution of a complex reflected ray must be added to the conventional real reflected ray in order to get a good high frequency approximation of the scattered field. The complex incident ray field is defined by the incident ray field in real space and by the analytical extension of the surface of the scatterer into a complex coordinate space, or equivalently by the analytical extension into complex space of the incident phase function along the scatterer surface. The reflected ray field is defined by imposing the continuity of the phase on the extended complex scatterer surface or by applying the law of reflection of complex rays. The reflected complex rays give non-specular contributions in real space. It is found that the complex reflection point is located near the reflection point on the concave side of the real obstacle surface (Fig. 8).

A similar effect exists when a concave-convex portion of a body is illuminated by a creeping wave. A nonspecular reflected creeping wave can therefore propagate in opposite direction and contribute to the monostatic return of the scatterer (Fig. 9). Obviously, a bistatic nonspecular return is also possible.

These phenomena are important in the calculation of radiation patterns of sources mounted on smooth bodies having concave portions. They are not present on a purely convex body since no complex stationary points exist on such a geometry.

Resolution of new canonical problems

In a paper by Idemen and Felsen [30] a new canonical problem was formulated and its exact asymptotic solution was given for the diffraction of a whispering gallery mode by the straight edge of a cylindrically curved sheet. The authors established the rigorous edge diffraction coefficient that describes the cylindrically radiating field diffracted by the edge and the launching coefficients for the creeping wave modes excited on the convex side and for the whispering gallery modes excited on the concave side of the cylindrical sheet. The procedure consists of two basic steps: first, the physical 2π interval for the cylindrical polar angle is replaced by a non-physical, infinitely extended angular coordinate space $-\infty < \varphi < \infty$ wherein the cylindrically curved thin sheet is modeled as a surface of radius $\rho = a$ extending from $\varphi = 0$ to $\varphi = -\infty$. This infinite angular space permits consideration of azimuthally nonperiodic wave phenomena in a cylindrical (ρ, φ) coordinate frame, and resolution of the problem by Fourier transform methods based on the Wiener-Hopf technique; secondly, the solution is evaluated asymptotically in a simple form interpretable in terms of GTD.

A similar procedure was applied by Albertsen and Christiansen [24,31] to the same geometry, when a creeping wave mode is incident on the edge from the convex side. The results obtained in both cases for the leading term of the asymptotic expansion of the field diffracted by the edge confirm the general GTD model which states that the combined surface-edge diffraction can be described as a sequence of separate edge and surface diffractions. Moreover, the leading term of the edge diffraction coefficient is given by the tangent plane approximation. For instance, the diffraction of a whispering gallery mode by the edge on a cylindrically curved thin sheet may be ascertained by modeling the edge geometry locally as that of the tangent half-plane and the incident model field in terms of the ray congruence striking the edge. The same results hold for a creeping mode when the diffracted space ray does not graze the sheet surface close to the edge. Another important consequence of the solutions of these canonical problems is the possibility to derive transition functions for a general curved wedge. Comments on this topic will be given in the next paragraph.

Derivation of transition functions for a curved wedge

By considering a line source illumination of a cylindrically curved sheet, and using adequate asymptotic expansions of the exact solution, Michaeli [32] established the expressions of the transition functions when the source point lies in region 4 and the observation point lies in region 2 or in region 3 (See Fig. 10). He also generalized the arguments of these functions to a general wedge angle and to a curved wedge. He found that the appropriate transition functions between regions 2 and 4 are the incomplete Airy function and its derivative [37] whereas between regions 3 and 4 the diffraction is described by Fock type surface reflection functions involving incomplete Airy functions [38,39]. These new special functions have not been tabulated so far. However efficient computer routines were developed [40].

Extension of GTD to metallic bodies with coatings

The general form of the asymptotic expansion of the field in the deep shadow of a smooth convex body given by (21) remains valid for a body with coatings. Indeed, it can be shown by solving rigorously the canonical problem of a metallic circular cylinder covered by an homogeneous coating and illuminated by a plane wave that the asymptotic sequence in the expansion (21) is compatible with the asymptotic expansion of the exact solution of this canonical problem. Moreover, a geometrical interpretation of the latter is given by the concept of creeping waves which follow the interface between the exterior medium and the coating. The generalization of GTD to bodies with coatings is therefore straightforward if we are able to define the boundary conditions on the exterior surface of a coating covering a general smooth metallic surface and derive the decay factor and the diffraction coefficients of the corresponding creeping waves.

Before analyzing this problem, it is instructive to come back to the canonical problem of a coated cylinder, illuminated by a plane wave. A first approach to this problem which is approximately valid when the refractive index of the coating is large, consists in replacing the coating by a constant surface impedance Z . By applying the Watson transformation to convert the classical series of radial modes into a series of angular modes propagating around the cylinder it is found that the asymptotic form of each angular mode may be interpreted as a creeping wave emanating from the shadow boundary and decaying exponentially with increasing distance from it along a path which is identical to that followed by a creeping wave on a perfectly conducting cylinder the radius of which is equal to the radius of the exterior surface of the coating. However, it appears also some fundamental differences with the perfectly conducting case, namely the decay factor as well as the behaviour of the field transverse to the ray depend strongly on the impedance Z . To show this, we denote by $K_p = k_0 \alpha_p$ the complex propagation constant of the p th mode where α_p is the local decay factor of that mode which is related to the total decay factor $\chi_p(Q')$ appearing in (21) by :

$$\chi_p(Q') = k_0^{-1/3} \int_Q^{Q'} \alpha_p(t') dt' \quad (28)$$

where the integration is performed along the path followed by the creeping ray from the shadow boundary Q to the point Q' where the ray leaves tangentially the surface of the cylinder.

For a perfectly conducting circular cylinder ($Z = 0$) with radius a , the values of K_p are given by $K_p = v_p/a$ where v_p are the roots of $H'_\nu(ka) = 0$ for TE waves and of $H'_\nu(ka) = 0$ for TM waves. They are located along the Stoke's line of the circular cylinder, functions in the first quadrant of the complex v -plane as shown in Fig. 11. For $Z \neq 0$, the values of K_p are given by the roots of the mixed boundary conditions :

$$H'_\nu(kb) + ik_0 \zeta H_\nu(kb) = 0 \quad (29)$$

where $\zeta = Z$ for TM waves and $\zeta = 1/Z$ for TE waves and where b is the radius of the outer surface of the coating. If a is the radius of the metallic cylinder and τ the thickness of the coating, we have $b = a + \tau$.

When $Z = i\alpha$ where α takes positive values from zero to infinity, the trajectories of K_p are found to be sinusoidal curves passing through the roots corresponding to $Z = 0$, and approaching tangentially the value $k_0 \sqrt{1+\alpha^2}$ on the real axis (Fig. 12). For complex values of Z , the trajectories may take one of the forms shown in Fig. 13 when $|Z|$ is varied from 0 to infinity [34]. It is seen that for small values of $\text{Re} Z$, the propagation constant of the first TM creeping mode leaves progressively the vicinity of the Stoke's line when $|Z|$ is augmented, and penetrates into the domain where the Hankel function may be represented for large $k_0 b$ by Debye's asymptotic expansion. It follows a modification of the physical characteristics of the wave which turns into a guided surface wave mode which has an exponential spatial behaviour transverse to its propagation direction whereas the transverse behaviour of a creeping wave is described by an Airy function which is the usual behaviour of a field near a regular caustic surface. These results remain roughly the same when the exact boundary conditions are applied on the exterior surface of the coated cylinder. Indeed, it can be shown that each angular mode satisfies a boundary condition of the form [35] :

$$H'_\nu(kb) + ik_0 \zeta(v) H_\nu(kb) = 0 \quad (30)$$

which is similar to (29), but now ζ is a function of v . As a consequence each TM mode (resp. TE mode) has a different impedance (resp. admittance) given by $Z(v_p)$ (resp. $Y(v_p)$) where v_p are the roots of (30).

The trajectories of the corresponding propagation constants in the complex v plane, when the thickness τ of the coating is varied, are shown on Fig. 14 for the first TM mode v_1 and the first TE mode v'_1 . We see that they have roughly the same form as in Fig. 12 for a lossless coating characterized by a real permittivity ϵ (perfect dielectric). However, the trajectory of the first TE mode does not pass through the root v'_0 corresponding to $Z = 0$ when the admittance becomes large but follows a curve which is approximately parallel to the trajectory of the first TM mode. This behaviour which appears also when the permittivity has a small imaginary value is a consequence of the fact that the equivalent admittance of a coating is not rigorously equal to the inverse of its equivalent impedance. When the imaginary part of the permittivity is augmented the trajectories of the roots of (30) leave the neighbourhood of the real v -axis and tend to become parallel to the Stoke's line as in Fig. 13 and finally transform into spirals as shown in Fig. 15 corresponding to a coated sphere. For a fictive boundary condition with $Y = \frac{1}{2}$ as in Fig. 12 and 13 the spirals transform into closed curves. Fig. 16 shows the trajectories of the creeping modes on a coated sphere when the thickness of a first layer is augmented from 0 to 1 cm and then a second layer with different characteristics is added and its thickness is augmented continuously.

When kb is large the Hankel function and its derivative in (30) can be replaced by their asymptotic expansions. In order to get results which can be easily transposed to a general smooth convex surface, we limit the number of parameters by choosing expansions which are appropriate in some regions of the complex v -plane but which are not uniform. In the region close to the Stoke's line limited by:

$$|(k_0 - k_p)b| \leq |k_p b|^{\frac{1}{3}} \quad (31)$$

An adequate expansion is given by the Olver expansion, in terms of Airy functions. If we insert in (30) the leading term of the expansion for $H_v(kb)$ and $H'_v(kb)$, we get the following equivalent equation:

$$\frac{A'_1(y)}{A_1(y)} = e^{-2i\pi/3} \gamma \quad (32)$$

where:

$$y = -2^{1/3} W v^{-1/3} e^{2i\pi/3}, \quad W = X - v, \quad X = k_0 b, \quad v = k_p b \quad (33)$$

and where γ is given by:

$$\gamma = \frac{\frac{1}{5v}(3W^2-1) + (i\zeta(v) + \Lambda)(1 - \frac{W}{5v^{2/3}})}{\frac{2^{1/3}}{v^{1/3}}(1 - \frac{4W}{5v^{2/3}}) - (i\zeta(v) + \Lambda) \frac{3 \times 2^{1/3} W^2}{10v^{2/3}}} \quad (34)$$

with $\Lambda = 0$ for a cylinder and $\Lambda = \frac{1}{2k_0 b}$ for a sphere.

Equations (32) to (34) depend only on the radius b of the outer surface of the coating and on the functions $\zeta(v)$ which are different for TM and TE modes. For a circular cylinder or a sphere, these functions are well known and their asymptotic expansions for τ/a small and $|k_1 a|$ large compared to $k_1 b$ where k_1 is the wave number in the medium constituted by the coating, can be easily derived by replacing the Hankel functions $H_v(k_1 a)$ and their derivatives appearing in the expression of $\zeta(v)$ by their Debye expansions. At first order, for a coating of thickness τ on a cylinder, they are given by:

$$Z(v) = -\frac{i\eta_1}{\eta} \psi(x_1, v) \operatorname{tg}[k_1 \tau \psi(x_1, v)] \quad \text{for TM modes} \quad (35)$$

$$Y(v) = \frac{i\eta}{\eta_1} \psi(x_1, v) \operatorname{cotg}[k_1 \tau \psi(x_1, v)] \quad \text{for TE modes} \quad (36)$$

where $\eta_1 = \sqrt{\frac{\mu_1}{\epsilon_1}}$ is the intrinsic impedance of the coating material and $\psi(x_1, v) = \sqrt{1 - \frac{v}{x_1^2}}$, with $x_1 = k_1 a$.

Fig. 14 shows that the error in the position of the roots of (30) by replacing the exact values of $\zeta(v)$ by approximate values given by (35) and (36) corresponds to a small shift along the trajectory which is most sensitive for small values of R_{0e} . Extensive numerical evaluations have shown that the expressions (35) and (36) give accurate results for the propagation constants for $\tau/a \leq 0.1$. For larger values of τ/a , additional terms must be added to (35) and (36). The extension of equation (32) with $\zeta(v)$ given by (35) and (36) to an arbitrary smooth surface covered by an homogeneous coating has been performed by Molinet [10]. He showed that for a torsionless geodesic on an otherwise arbitrary smooth surface, the local value of γ is still given by (34) but with Λ replaced by $\frac{1}{2k_0 \rho_{tn}}$ where ρ_{tn}

is the radius of curvature of the surface in the direction of the binormal to the geodesic followed by the creeping wave. In addition, the radius b appearing in γ and in relations (33) and the radius a in (35) and (36) must respectively be replaced by ρ_g and $\rho_g - \tau$ where ρ_g is the local radius of curvature of the geodesic and τ the local thickness of the coating. This result has been obtained by developing the solution of (32) around the zeros of $A'_1(y)$ (resp. $A_1(y)$) with γ given by (34) where Λ is supposed unknown, and with $\zeta(v) = 0$ (resp. $\zeta(v) = \infty$) and by comparing the terms of this development with those found by Voltmer [36] for an arbitrary smooth surface.

Extension of the preceding formulas to a geodesic with torsion is straightforward by following the method outlined in [11] by Levy and Keller.

In order to calculate the field associated with each creeping ray, we need to extract from the canonical problem of the cylinder or the sphere, expressions for the diffraction coefficients. Levy and Keller [11] showed that the dyadic diffraction coefficient corresponding to the p th angular mode, may be expressed by:

$$D_p(Q, Q') = \hat{n}(Q)\hat{n}(Q')D_p^{TM}(Q)D_p^{TM}(Q') + \hat{b}(Q)\hat{b}(Q')D_p^{TE}(Q)D_p^{TE}(Q') \quad (37)$$

where Q and Q' are the initial and final points of the geodesic followed by the creeping wave, \hat{n} is the unit vectors along the normal and binormal to the geodesic and where $D_p^{TM, TE}$ are the surface diffraction coefficients for TM and TE waves respectively. It follows from (37) that the normal and tangential components of the incident field at Q travel independently along the geodesic and that no coupling between these two components occurs, at least in the leading term of the asymptotic expansion. This result which has first been established for a perfectly conducting surface remains valid for a coated surface under the hypothesis adopted above ($|k_1 a|$ large and τ/a small). This conclusion holds only approximately. In fact for oblique incidence a coupled transcendental equation is obtained which shows that the TE and TM components are coupled.

From the exact solutions of the coated cylinder and sphere in terms of angular modes, it is found by replacing the Hankel functions and their derivatives with respect to v by their appropriate asymptotic expansions using Olver formula, that the surface diffraction coefficients in (37) are given by :

$$D_p(Q) = \frac{2e^{\frac{3i\pi}{8}}}{\sqrt{Q_n}} \left(\frac{2\pi}{k_0} \right)^{1/4} \quad (38)$$

$$Q_n = \frac{2^{10/3} e^{2i\pi/3}}{v^{5/3}} \pi k b [A_1^2(y) - y A_1^2(y)] + \frac{8\pi k b}{v^2} \left[-\frac{47}{30} y^2 A_1^2(y) + \frac{19}{15} y A_1^2(y) - \frac{3}{5} A_1(y) A_1'(y) \right] \quad (39)$$

where y is defined in (33) with $x = k_0 \rho_g$. The general expressions (38) and (39) are the same for the TM and TE modes. The diffraction coefficients D_p^{TM} and D_p^{TE} are derived from these expressions by replacing v by the roots of (32) for TM and TE modes respectively.

A convenient and widely used notation for the total field diffracted by all the modes at an observation point P is (13, 14) :

$$\vec{E}^d(P) = \vec{E}^i(Q) \cdot [\hat{A}(Q)\hat{A}(Q')F + \hat{b}(Q)\hat{b}(Q')G] \frac{e^{jks}}{s(\rho+s)} e^{jks} \quad (40)$$

in which :

$$G = e^{ikt} \left(\frac{dn(Q')}{dn(Q)} \right)^{-1/2} \sum_{p=1}^{\infty} D_p^{TM}(Q) D_p^{TM}(Q') e^{-\int_Q^{Q'} \alpha_p^{TM}(t') dt'} \quad (41)$$

and G has the same form as F except that the super script TM is replaced by TE. The factor $\left[\frac{dn(Q')}{dn(Q)} \right]^{-1/2}$ takes into account the amplitude variation resulting from the divergence of the creeping wave. On a cylindrical surface, illuminated by a plane wave, this factor is equal to unity. The local decay factor α_p is equal to the imaginary part of the propagation constants K_p of the creeping modes. The asymptotic solution above is not valid in the transition region 1 of Fig (6) where the behaviour of the field is described by a Fock type transition function. For a boundary surface characterized by a constant impedance Z , this function has the general form :

$$P(\xi) = \frac{e^{i\pi/4}}{\sqrt{\pi}} \int_{-\infty}^{+\infty} \frac{V'(t) - im \zeta_1(t)}{W_1(t) - im \zeta_2 W_2(t)} e^{i\xi t} dt \quad (42)$$

where $V(t)$ and $W_2(t)$ are Airy functions and $V'(t)$ and $W_1'(t)$ their derivatives with respect to the argument, defined in [39] and where $\xi = \frac{z}{\rho}$ for a TM wave and $\xi = \frac{\eta}{\rho}$ for a TE wave. The local geometry of the surface enters through the parameter m . For a sphere of radius a , one obtains $m = (\frac{ka}{2})^{1/3}$. For $z = 0$, (42) reduces to the Fock type surface reflection functions tabulated by Logan [38]. For a coated surface, $\zeta = \zeta(t)$ is a function of the integration variable t , which is known for a cylinder or a sphere. Indeed, ζ is given in terms of the Bessel and Hankel functions of indice v and in deriving (42), the transformation $v = k_0 b + mt$ has been used. However, since the expression of ζ is rather complicated, the direct numerical computation of the integral in (42) is extremely time consuming. Efficient subroutines have been recently developed at MOTHEM for the calculation of this integral. Basically, the method consists in replacing ζ by a constant value $\zeta(\tau_1)$ where τ_1 is related to the first root of the characteristic equations (30) or (32) for creeping modes. Fig (17) and (18) show results obtained for a coated circular cylinder. A different algorithm has been used in each region. In the transition region the diffracted field has been calculated by the numerical evaluation of (42) with our optimized subroutine. Away from the transition region, we have applied GO in the illuminated region and the creeping wave formalism in the deep shadow region. We see that the numerical results are continuous through the boundaries of these three regions which is a check of the validity of our algorithm. Progress has also been made in the resolution of the diffraction of a plane wave by an impedance straight wedge by extending the solution obtained for normal incidence by Maliuzhinets [41]. The diffraction problem of an electromagnetic plane wave by a half-plane with equal face impedances has been solved for oblique incidence by Senior [42] and with different face impedances and oblique incidence by Bucci and Franseschetti [43]. The electromagnetic diffraction at oblique incidence of a plane wave by a right-angled wedge one face of which is imperfectly conducting and the other metallic, has been solved by Vaccaro [44] who extended Maliuzhinets generalized reflection method. Uniform solutions through the shadow boundaries of the direct and reflected GO fields have also been derived. An extensive literature on this subject appeared in the last five years [16]. However, the diffraction problem of an electromagnetic plane wave obliquely incident on an impedance wedge with an arbitrary interior angle still remains unsolved. Present investigations concern the extension of Maliuzhinets solution to surfaces with coatings, by applying generalized approximate boundary conditions or by introducing exact reflection coefficients in the plane wave spectrum of the diffracted field. Results obtained with the last method are shown on fig. 19 and 20.

Conception of a general computer program founded on GTD/UTD

It follows from the presentation in the preceding chapters, that GTD and its uniform versions, constitute a multiplicity of asymptotic solutions, each of them having a limited domain of validity. Indeed, even those solutions which have been called "uniform" in our former description are in fact only partially uniform since they do not take into account

all the parameters of the problem. Moreover, it happens frequently that a partially uniform solution be less accurate in some regions of space than a uniform solution or that it needs too much computer time for its evaluation. The first step in constructing a general computer program consists therefore in selecting in each region of space and for each diffraction process, the most efficient asymptotic solution available for the computation of the diffracted field.

For example, if we consider the diffraction of a local plane wave by a curved wedge, the space outside the wedge angle must be divided into six different regions labelled from 1 to 6 (Fig. 5). We see that regions 4, 5 and 6 are independent of the direction of propagation of the incident ray, whereas regions 2 and 3 depend on that direction. Furthermore when the incoming ray grazes one of the faces of the wedge, regions 2 and 3 or regions 2 and 4 overlap. Hence, we count five or four different solutions for non grazing incidence depending whether we use a partially uniform solution through regions 1, 2 and 3 or not. For grazing incidence, two other solutions must be added, giving a total of 6 or 7 different solutions, for an incoming plane wave. Moreover when the incident field is a transition region field three complementary solutions are needed. Similar conclusions hold for other diffraction processes. Consequently, the choice of the appropriate asymptotic solution for each couple of incoming and diffracted rays leads to a management problem which can only be solved by a computer. For each diffraction process the solution of this problem depends on the direction of the incoming and diffracted rays, and on a set of geometrical parameters which enter in the evaluation of the angular extension of the various transition regions such as the principal directions and the radii of curvature of the incoming wave and of the surfaces in the vicinity of the diffraction points. A separate subroutine can therefore be conceived for each diffraction process, and the whole GTD/UTD can be stocked in the form of a library of subroutines managed by a main program. The field computation is performed by following the trajectory of a ray and by calling the appropriate subroutine at its points of interaction with the scatterer. This was the basic idea underlying the construction by MOTHE-SIM of a general computer library for GTD applications called PROMETHEE. Two versions of this library have been developed: a version V1 for perfectly conducting bodies and a version V2 for coated metallic bodies. In V2 diffraction processes involving surface rays are also included. It is worth while mentioning that PROMETHEE is completely independent of the geometrical modelling of the target and of the corresponding ray searching techniques. It can therefore be coupled to any such techniques. Once a ray trajectory has been found with an arbitrary number of interactions with the target, the corresponding parameters including those characterizing the surfaces close to each diffraction point, are transferred to PROMETHEE which calculates the corresponding field. At present, PROMETHEE treats the following interactions on a scatterer with or without coatings: reflection, double reflection at grazing angles, edge-diffraction, double edge diffraction in transition regions, creeping waves, edge-diffracted creeping waves, diffraction by an impedance discontinuity in a regular surface, edge diffracted surface waves, diffraction of a surface wave or a creeping wave by an impedance discontinuity, diffraction by a curvature discontinuity, tip diffraction, corner diffraction, tip excited or diffracted surface waves.

Great strides have also been made in modelling of targets and in ray searching techniques. The shapes to which GTD is applied are of growing complexity and the need for general computer codes which manage automatically the geometrical modelling of a target once a set of data is introduced with the aid of an interactive procedure, becomes more and more pronounced. We illustrate here briefly a technique developed by THOMSON-CSF which consists in modelling a given target by simple shapes: like plates, circular and elliptic finite cylinders, circular and elliptic finite cones or frustums, truncated prolate ellipsoids, and more general frustums generated by straight lines which lean on two planar curves (circle, ellipse, ...). Fig. 21 shows some typical targets which have been modelled with this technique. For the fighter, on figure 21(a), about 40 elementary surfaces have been needed in the geometrical model.

Ray searching techniques associated with these geometrical modelling have also been developed by THOMSON-CSF. The choice of analytical surfaces greatly simplifies the determination of the ray trajectories which is performed by Newton and Min-Max algorithms. At present, the number of interactions is limited to three and the number of diffraction processes other than reflections is limited to one. For example triply reflected rays and doubly reflected rays which undergo in addition an edge diffraction or a surface diffraction (creeping ray) belong to the set of rays which are determined. An interactive simulator called SARGASSES for antennas on structures such as an aircraft or a spacecraft, using the techniques described above for the geometrical modelling and ray searching, has been created by THOMSON-CSF. It is coupled with PROMETHEE for the calculation of the field associated with each ray. The software is designed around three principal functions:

- data description (geometry of the structures, antenna characteristics).
- digital processing (search for ray trajectory and computation of the associated electromagnetic fields).
- interpretation of the results.

Fig. (22) and (23) show the ray trajectories obtained with SARGASSES between a source located in the vicinity of a finite cylinder and a frustum and an observation point at infinity, limited to one and two interactions respectively. A more realistic example of ray tracing corresponding to the radiation of a source on the surface of the spacecraft HERMES is shown on fig. (24).

4. SOME APPLICATIONS

We present two types of applications. The first concerns the radiation of antennas mounted on the surface of an aircraft or a spacecraft. The second concerns the coupling between reflector antennas located on a ship. In both situations, the interaction with the vehicle supporting the antennas is taken into account.

Radiation of antennas mounted on an aircraft or a spacecraft

We are primarily interested in the control of the deformation of the radiation patterns.

Our first example concerns antennas mounted on the lee-board of a MIRAGE 2000 fighter. The influence of the aircraft structure on the radiation pattern of the antenna was carried out using SARGASSES software. Fig. 25 a,b shows the graphical output of the ray searching algorithm which has been limited to single and double interactions. It is seen that seven to ten different rays are found in an observation direction close to the axis of the aircraft. Fig. 25 c gives the volume directivity pattern of the same antenna taking into account the interactions with the aircraft and Fig. 25 d shows an azimuthal conical pattern around a vertical axis at an angle of 10° away from this axis. The field computations have been performed with PROMETHEE version 1, the structures being supposed perfectly conducting.

Our second example corresponds to a source located on a support fixed under the right wing of the same fighter (Fig. 26 a). The results obtained with SARGASSES for the vertical polarization at a frequency in the X band are shown on fig. 26 b. They compare favourably with those obtained by measurements accomplished with some antennas mounted on a real scale model of the fighter shown on Fig. 26 c.

Our third application of GTD/UTD concerns the analysis of the radiation pattern of 4 antennas mounted on the fuselage of the HERMES spacecraft as shown on Fig. 27a,b. The influence of the spacecraft antenna radiation was again carried out using SARGASSES. Fig. 27c and 27d show respectively the three-dimensional directivity pattern of the four antennas with and without the interactions of the radiated field with the structures of the spacecraft. It is found that the influence of the spacecraft is principally a shadowing effect. Away from the shadow region of the direct rays, the interactions with the spacecraft have only a small effect on the directivity pattern.

Coupling between reflector antennas located on a ship

The coupling between reflector antennas on a ship takes generally place in the near field region (Fresnel zone) of the radiation patterns of the antennas. However, if we suppose that direct illumination of the receiving antenna by the main lobe of the transmitting antenna has been avoided, the coupling problem of two reflector antennas in the near field can be reduced to a coupling problem in the far field by treating the reflectors as obstacles which diffract the field radiated by the primary sources. The initial problem becomes then a problem which can be solved by GTD. Each primary source is defined by its phase center O and by its normalized vectorial radiation pattern $f(\hat{u})$ where $\hat{u} = \hat{u}(\theta, \phi)$ is the radiation direction and (θ, ϕ) are spherical polar coordinates with respect to an appropriate direction OZ . It can be shown that the power P_r received at a given frequency F by a receiving reflector antenna located in the vicinity of a transmitting reflector antenna is given by :

$$P_r = \frac{\lambda^2}{16\pi^2} P_t G_t(o) G_r(o) \sum_{i=1}^N \left| \bar{D}_i \cdot \hat{f}_t(\hat{u}_i^t) \cdot \hat{f}_r(-\hat{u}_i^r) \right|^2 \quad (43)$$

where the symbols have the following meaning :

- P_t : maximum power of the transmitter at the frequency F .
- λ : wavelength = $\frac{c}{F}$ where c is the speed of light.
- $G_t(o)$: maximum gain of the primary source of the transmitting antenna.
- $G_r(o)$: maximum gain of the primary source of the receiving antenna.
- N : total number of ray trajectories from the phase center of the source of the transmitting antenna to the phase center of the source of the receiving antenna.
- \hat{u}_i^t, \hat{u}_i^r : unitary vectors in the direction of emission of ray i at the transmitter and in the direction of arrival of ray i at the receiver.
- \bar{D}_i : dyadic diffraction coefficients taking account of all the diffraction processes encountered by ray i between the transmitting and receiving primary sources.

The number of edge-diffractions has been limited to two (order k_0^{-1}) since the term of order $k_0^{-3/2}$ is not available for a curved edge (see chap. 2). It can be shown that nine classes of rays verifying this condition can exist between two Cassegrain reflector antennas if we neglect doubly diffracted rays on the same reflector. Between two parabolic reflectors and between a parabolic reflector and a Cassegrain antenna, respectively four and six different classes of rays can exist. It is found that when the interactions with the ship are not taken into account, each class of rays is constituted by a maximum of four rays except for the class of direct rays emanating from the primary source (spill-over) and joining directly the phase center of the primary receiving antenna which contains only one ray. Owing to the fact that the order of the asymptotic series for edge diffraction is limited to k_0^{-1} , the interactions with the ship must be limited to reflections except for the rays which do not intercept the reflectors (direct rays), for which single edge-diffraction with the ship structures must be added.

A computer program has been developed at MOTHESIM for the resolution of this coupling problem. The ship is modelled by polygonal flat plates. Ray searching algorithms determine all the classes of rays described above. The field along a ray is calculated with PROMETHEE.

REFERENCES

8-15

- [1] A. SOMMERFELD and J. RUNGE, "Anwendung der Vektorrechnung auf die Grundlagen der geometrischen Optik", Ann. Phys., Vol. 35, pp. 277-298, 1911.
- [2] H. POINCARÉ, 1892, "New methods of celestial mechanics", Vol. 1-3 (English translation), NASA TTS-450, 1967. 10,56.
- [3] B.C. IGNATOWSKY, Trans. State Opt. Inst. (Petrograd), Vol. 1, n° 3, 1919.
- [4] F.G. FRIEDLANDER, "Geometrical Optics and Maxwell's equations", Proc. Cambridge Phil. Soc., Vol. 43, p. 284, 1947.
- [5] J.B. KELLER, "Diffraction by an aperture", J. Appl. Phys., Vol. 28, pp. 426-444, 1957.
- [6] J.B. KELLER, "A geometrical theory of diffraction", Proceedings of Symposia in Applied Mathematics, Vol. VIII: Calculus of Variation and its Application, pp. 27-52, Mc. Graw-Hill, Book Co., New-York 1958.
- [7] J.B. KELLER, "Diffraction by a convex cylinder", IRE Trans. Ant. Prop., Vol. AP-4, pp. 312-321, July 1956.
- [8] F.G. FRIEDLANDER and J.B. KELLER, "Asymptotic expansions of solutions of $(\nabla^2 + k^2)u = 0$ ", Comm. Pure and Applied Math., Vol. 8, pp. 387-394, 1955.
- [9] R.N. BUCHAL and J.B. KELLER, "Boundary layer problems in diffraction theory", Comm. Pure and Applied Math., Vol. 13, pp. 85-114, 1960.
- [10] F. MOLINET, "Geometrical Theory of Diffraction (GTD), Part I: Foundations of the theory, IEEE Ant. Prop. Newsletter, pp. 6-17, August 1987, Part II: Extensions and future trends of the theory, IEEE Ant. Prop. Newsletter, pp. 5-16, October 1987.
- [11] B.R. LEVY and J.B. KELLER, "Diffraction by a smooth object", Comm. Pure and Appl. Math., Vol. 12, pp. 159-209, February 1959.
- [12] R.G. KOUYOUMJIAN and P.H. PATHAK, "A geometrical theory of diffraction for an edge in a perfectly conducting surface", Proc. IEEE, Vol. 62, pp. 1448-1461, 1974.
- [13] P.H. PATHAK and R.G. KOUYOUMJIAN, "An Analysis of the Radiation from Apertures in curved Surfaces by the Geometrical Theory of Diffraction", Proc. IEEE, Vol. 62, N° 11, pp. 1409-1447, Nov. 1974.
- [14] P.H. PATHAK, "Uniform Geometrical Theory of Diffraction" in "Theoretical Aspects of Target Classification" AGARD L.S. N° 152, June-July 1987, pp. 2.1 - 2.20.
- [15] R. LEWIS and J. BOERSMA, "Uniform Asymptotic Theory of edge diffraction", J. Math. Phys., Vol. 10, N° 12, pp. 2291-2305, Déc. 1969.
- [16] D.S. AHLUWALIA, "Uniform asymptotic theory of diffraction by the edge of a three-dimensional body", SIAM J. Appl. Math., Vol. 18, pp. 287-301, 1970.
- [17] S.W. LEE and G.A. DESCHAMPS, "A Uniform Asymptotic Theory of electromagnetic diffraction by a curved wedge", IEEE Trans. Ant. Prop., Vol. AP-24, pp. 25-34, Jan. 1976.
- [18] W. PAULI, "On asymptotic series for functions in the theory of diffraction of light", Phys. Rev., Vol. 54, pp. 924-931, 1938.
- [19] P.C. CLEMMOW, "Some extensions to the method of integration by steepest descents", Quart. J. Mech. Appl. Math., 3, pp. 241-256, 1950.
- [20] F. OBERHETTINGER, "On asymptotic series for functions occurring in the theory of diffraction of waves by wedges", J. Math. Phys., Vol. 34, pp. 245-255, 1956.
- [21] VAN DER WAERDEN, "On the method of saddle points", Appl. Sci. Rsch., B2, pp. 33-45, 1951.
- [22] Y.H. HWANG and R.G. KOUYOUMJIAN, "A dyadic diffraction coefficient for an electromagnetic wave which is rapidly varying at an edge". USNC-URSI Annual Meeting, Boulder, Colorado, 1974. For explicit expressions see also: R.G. KOUYOUMJIAN, P.H. PATHAK and D. BURNSIDE, "A uniform GTD for the diffraction by edges, vertices and convex surfaces", in Theoretical Methods for Determining the Interaction of Electromagnetic waves, with structures, J.K. SKWIRZYNSKI, ed., Netherlands, Sijthoff and Nordhoff, 1981.
- [23] J. BOERSMA and Y. RHAMAT-SAMII, "Comparison of two leading uniform theories of edge diffraction with the exact uniform asymptotic solution", Radio Sci., Vol 15, N° 6, pp. 1179-1194, 1980.
- [24] N.C. ALBERTSEN and J. CHRISTIANSEN, "Hybrid diffraction coefficients for first and second order discontinuities of two dimensional scatterers", SIAM J. Appl. Math., 34,

- pp. 398-414, 1978.
- [25] **M. IDEMEN**, " Diffraction by an obliquely incident high-frequency wave by a cylindrically curved sheet ", IEEE Trans. Ant. Prop., AP-34, N°2, pp. 181-187, Feb. 1986.
 - [26] **D. LUDWIG**, " Uniform asymptotic expansion at a caustic ", Comm. Pure Appl. Math., Vol. 19, pp. 215-250, 1966.
 - [27] **P. Ya. UFIMTSEV**, " Method of edge waves in the Physical Theory of Diffraction " Air Force System Command, Foreign Tech. Dir., Doc. ID N° FTD-HC-23-259-71 released for public distribution, Sept. 7, 1971 (Translation from the Russian version published by Soviet Publication House, Moscow, 1962).
 - [28] **A. MICHAELI**, " Elimination of infinities in equivalent edge currents " Part I : " Fringe current components ", Part II : " Physical optics components ", IEEE Trans. Ant. Prop., Vol. AP-34, pp. 912-918 and AP-34, pp. 1034-1037, Aug. 1986.
 - [29] **H. IKUNO** and **L.B. FELSEN**, " Complex ray interpretation of reflection from concave-convex surfaces ", and " Complex rays in transient scattering from smooth targets with inflection points ", IEEE Trans. Ant. Prop., Vol. AP-36, pp. 1260-1271 and 1272-1280, Sept. 1988.
 - [30] **M. IDEMEN** and **L.B. FELSEN**, " Diffraction of a whispering gallery mode by the edge of a thin concave cylindrically curved surface ", IEEE Trans. Ant. Prop., Vol. AP-29, pp. 571-579, 1981.
 - [31] **N.C. ALBERTSEN**, " Diffraction of creeping waves ", Res. Rep. LD 24, Technical University of Denmark, Electromagnetic Institute, Lyngby, Denmark 1974.
 - [32] **A. MICHAELI**, " Transition functions for high frequency diffraction by a curved perfectly conducting wedge. Part I : Canonical solution for a curved sheet, Part II : A partially uniform solution for a general wedge angle, Part III : Extension to overlapping transition regions ", submitted to IEEE Trans. Ant. Prop.
 - [33] **G.L. JAMES**, " Geometrical Theory of Diffraction for Electromagnetic Waves ", Hertfordshire, England, Peter Peregrinus, 1976.
 - [34] **P.R. BRAZIER-SMITH** and **J.F. SCOTT**, " Sound propagation over curved boundary surfaces ", Journal of Sound and Vibration 95(2), pp. 232-235, 1984.
 - [35] **R. PAKNYS** and **N. WANG**, " Creeping wave propagation constants and modal impedance for a dielectric coated cylinder " IEEE Trans. Ant. Prop., Vol. AP-34, N° 5, pp. 674-680, May 1986, - " Excitation of creeping waves on a circular cylinder with a thick dielectric coating ", IEEE Trans. Ant. Prop., Vol. AP-35, N° 12, pp. 1487-1489, Dec. 1987.
 - [36] **D.R. VOLTMER**, " Diffraction by doubly curved convex surfaces ", Ph. D. Dissertation, The Ohio State University, Columbus, Ohio, 1970.
 - [37] **L. LEVEY** and **L.B. FELSEN**, " On incomplete Airy functions and their application to diffraction problems ", Radio Science, Vol. 4, N° 10, pp. 959-969, Oct. 1969.
 - [38] **N.A. LOGAN**, " General Research in diffraction theory ", Vol I, LMSD-288-87 and Vol II, LMSD-288-88, Missiles and Space Division, Lockheed Aircraft Corporation, 1959.
 - [39] **J.J. BOWMAN**, **T.B.A. SENIOR** and **P.L.E. USLENGHI**, Eds., " Electromagnetic and Acoustic Scattering by Simple Shapes ", Amsterdam, the Netherlands : North Holland Publ., 1969.
 - [40] **T. CWIK**, " On the efficient calculation of the incomplete Airy function with application to edge diffraction ", submitted to Radio Science.
 - [41] **G.D. MALIUZHINETS**, " Excitation, reflection and emission of surface waves from a wedge with given face impedances " Sov. Phys. Dokl., 3, pp. 752-755, 1958.
 - [42] **T.B.A. SENIOR**, " Diffraction by an imperfectly conducting half-plane at oblique incidence ", Appl. Sci. Res., Sect. B, 8, pp. 35-61, 1960.
 - [43] **O.M. BUCCI** and **G. FRANCESCHETTI**, " Electromagnetic scattering by an edge with two face impedances ", Radio Science, Vol. 11, N°1, pp. 49-59, Jan.
 - [44] **V.G. VACCARO**, " Electromagnetic diffraction from a right-angled wedge with soft conditions on one surface ", Optica Acta, Vol. 28, N° 3, pp. 293-311, 1981.

ACKNOWLEDGEMENTS

The author wishes to thank the company THOMSON-CSF - Division RCM who provided the results on the MIRAGE 2000 Fighter and the company ALCATEL-ESPACE (ATES) who provided the results on the spacecraft HERMES. Thanks are also addressed to the military administration DRET who supported the development of PROMETHEE.

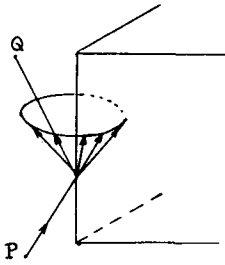


Fig. 1 : Edge diffraction

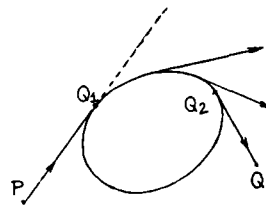


Fig. 2 : Creeping rays

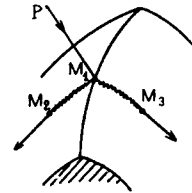


Fig. 3 : Edge excited creeping rays

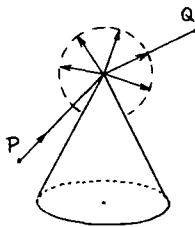


Fig. 4 : Tip diffraction

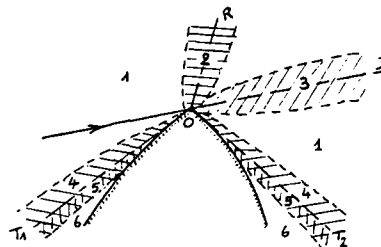


Fig. 5 : Domains of validity of various uniform edge solutions

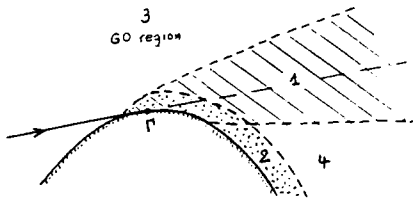


Fig. 6 : Domains of validity of various uniform solutions for surface diffraction

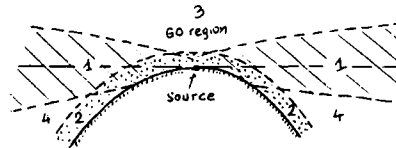


Fig. 7 : Transition regions for a source on a smooth surface

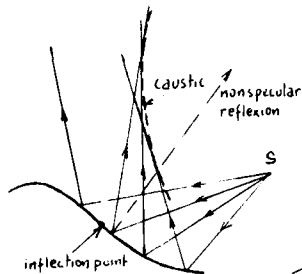


Fig. 8 : Reflection from a concave-convex boundary



Fig. 9 : Nonspecular reflected creeping wave

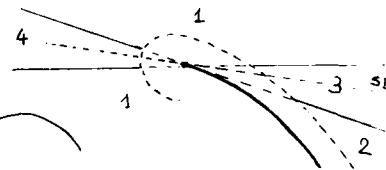


Fig. 10 : Transition regions of a curved sheet

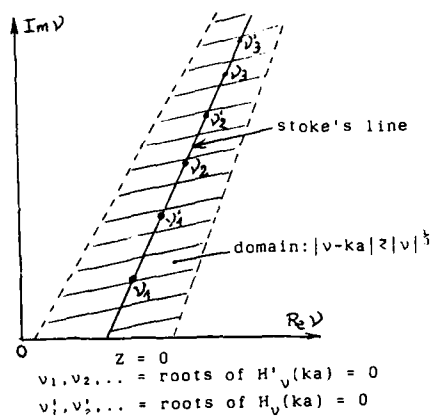


Fig. 11 : Propagation constants for a perfectly conducting cylinder

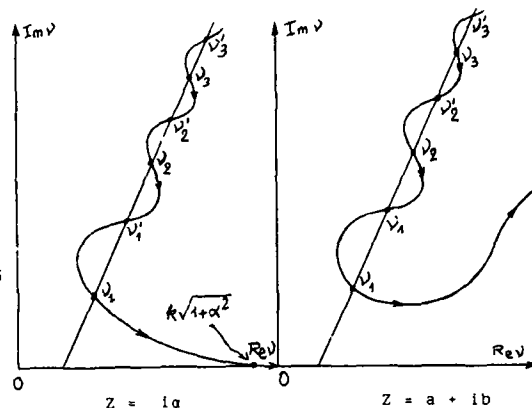


Fig. 12

Fig. 13

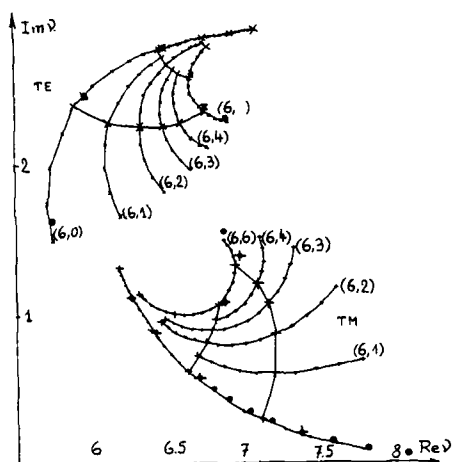
Trajectories of the azimuthal propagation constants when $|Z|$ varies

Fig. 14 : Coated cylinder

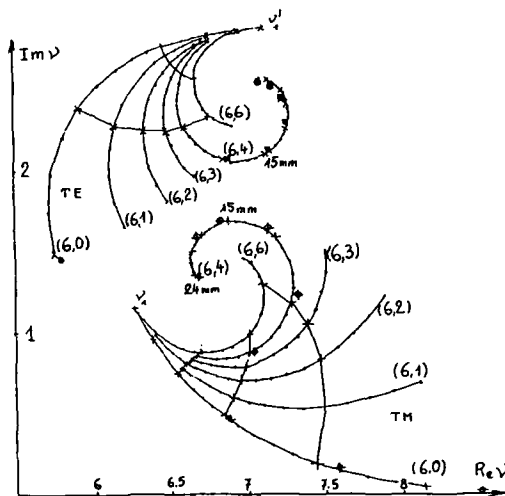


Fig. 15 : Coated sphere

Trajectories in the complex plane, of the azimuthal propagation constants of the first TM and the first TE creeping wave modes as a function of the coating thickness

Relative permittivity of the coating :

$\epsilon_1 = 6 + i8 = (6,8)$, $a = \text{variable}$

$b = 5,5$, $b = \text{exterior radius of the sphere}$

$\tau = \text{thickness of the coating, varies from 0 to 12 mm}$

(from 0 to 24 mm for $\epsilon_1 = 6,4$ on the coated sphere)

x—x—x exact values for $Z(\nu)$, $Y(\nu)$

• • • approximate values given by equations (35) and (36).

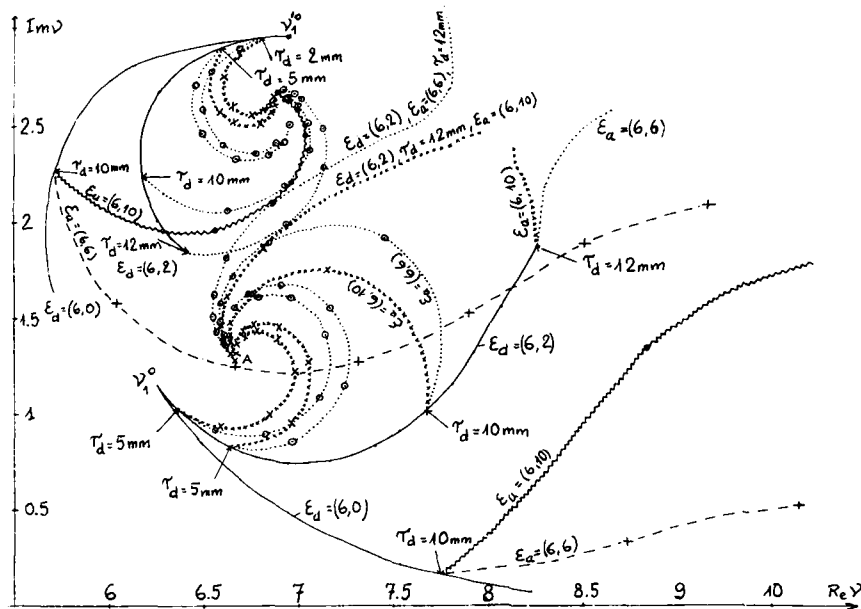


Fig. 16 : Coated sphere (2 layers)
Legend as in Fig. 14, 15

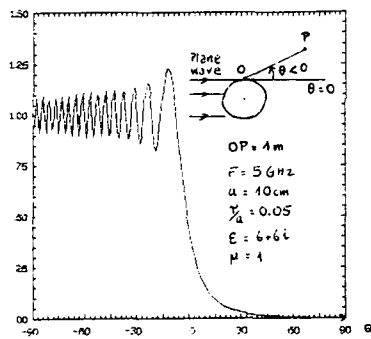


Fig. 17 : Total field diffracted by a coated cylinder : TE case

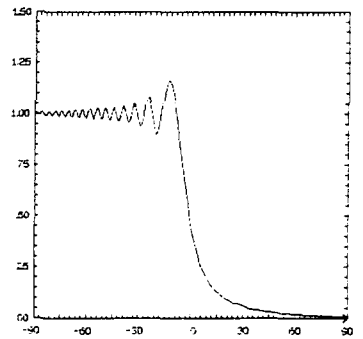


Fig. 18 : Total field diffracted by a coated cylinder : TM case

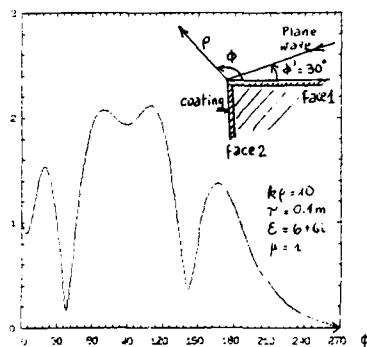


Fig. 19 : Total field diffracted by a coated right-angled wedge : TE case

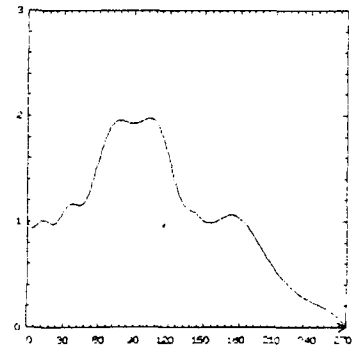


Fig. 20 : Total field diffracted by a coated right-angled wedge : TM case

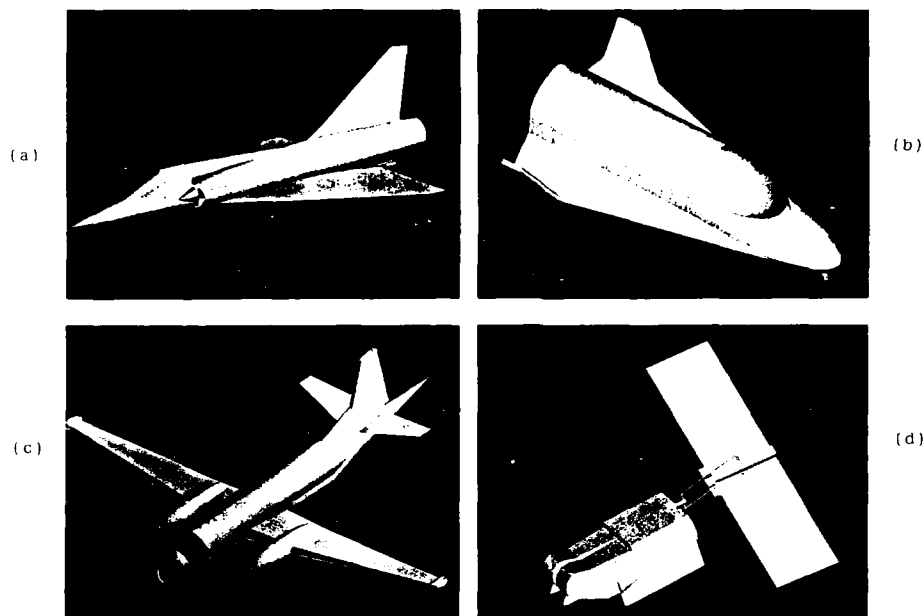


Fig. 21 : Geometrical modelling of targets with simple shapes (SARGASSES)

CYLINDRE ET CONE
Fichier : SIMPLES INTERACTIONS
Position source
X= 0.00 Y= 5.00 Z= 4.00
Direction d'observation
THETA= 30.0 PHI= 85.0

— Rayon direct
— Rayon reflechi
— Rayon diffracte
— Rayon rampant

SARGASSES

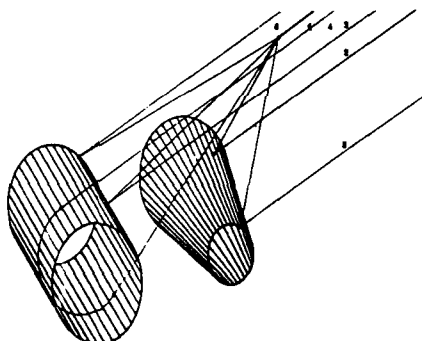


Fig. 22 : Ray searching : single interactions

CYLINDRE ET CONE
Fichier : DOUBLES INTERACTIONS
Position source
X= 0.00 Y= 5.00 Z= 4.00
Direction d'observation
THETA= 30.0 PHI= 85.0

— Rayon direct
— Rayon reflechi
— Rayon diffracte
— Rayon rampant

SARGASSES

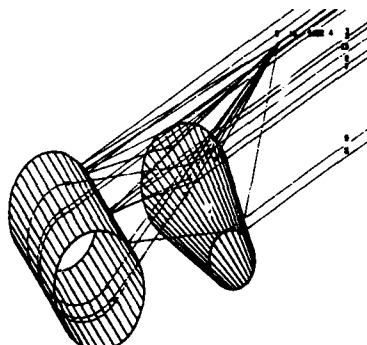


Fig. 23 : Ray searching : double interactions

HERMES
 Fichier : SARGASSES
 Position source : 1
 X= 0.46 Y= 1.68 Z= 0.00
 Direction d'observation
 THETA= 45.0 PHI= 60.0

—x— Rayon direct
 —xx— Rayon reflechi
 —xxx— Rayon diffracte
 — Rayon rampant

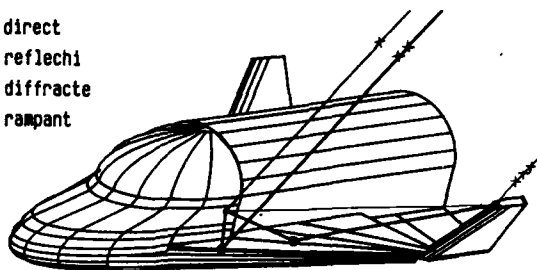
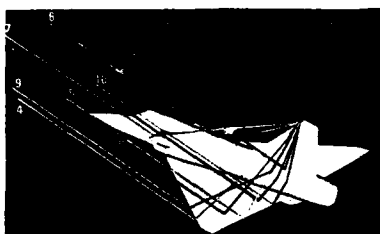
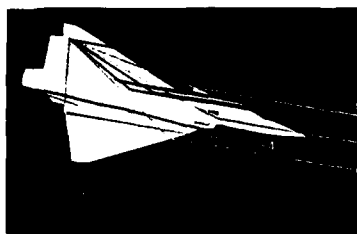


Fig. 23 : Rays emanating from a source on the spacecraft HERMES

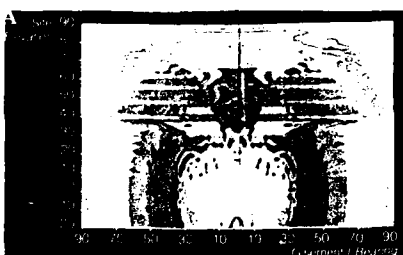


(a)

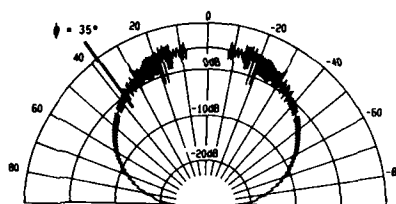


(b)

Graphical output of the ray searching algorithm (SARGASSES)



(c) Volume directivity pattern



(d) Azimuthal conical pattern (F=10 GHz, E_θ)

Fig. 25 : Antenna on the lee-board of a MIRAGE 2000 fighter

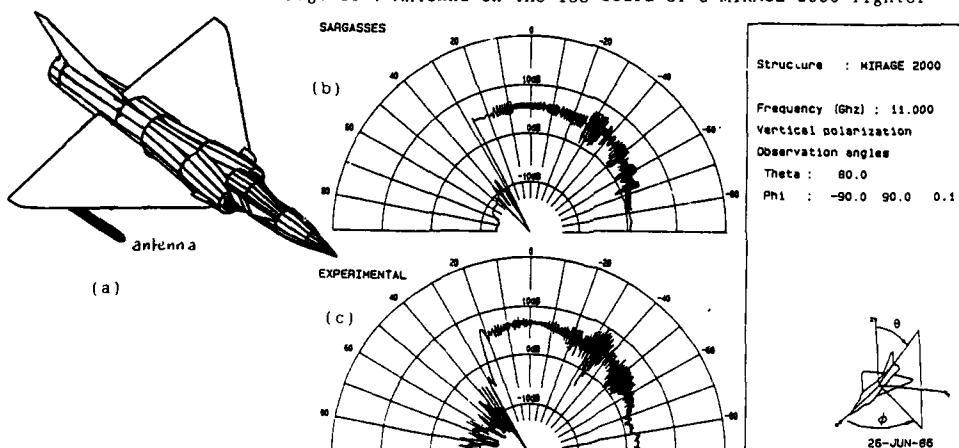
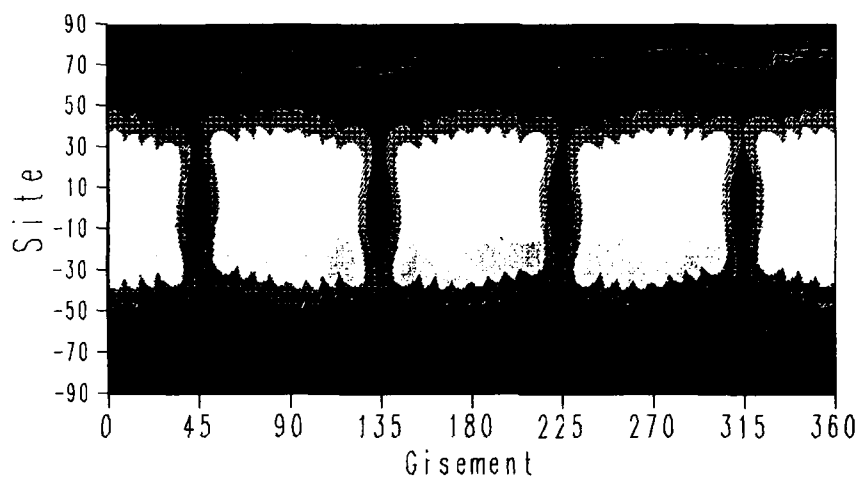
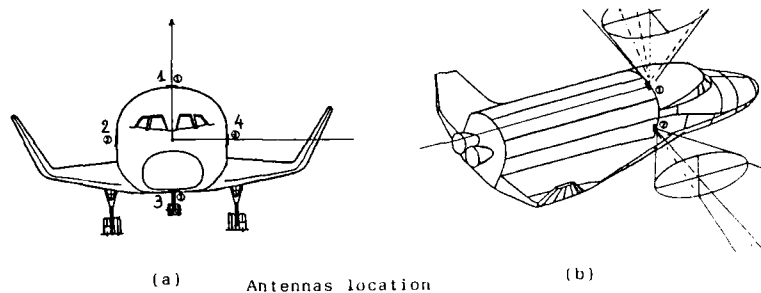
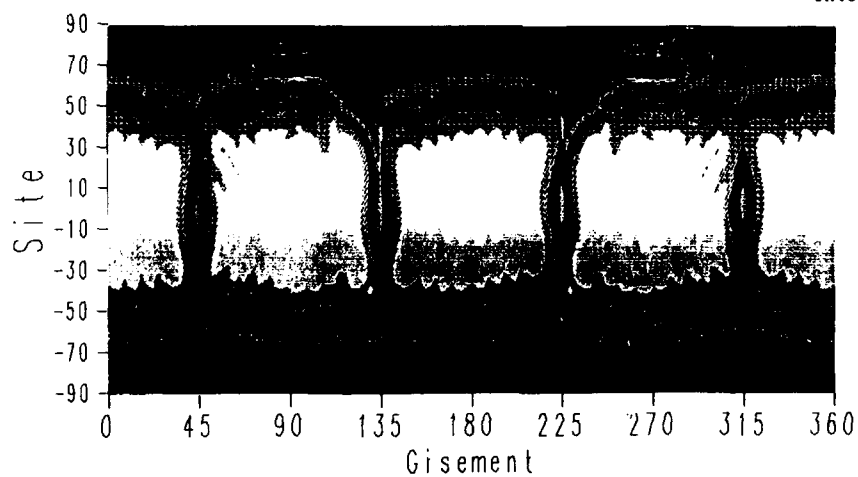


Fig. 26 : Antenna mounted on a support under the wing of a MIRAGE 2000



(c) Curves of constant levels of the three-dimensional directivity pattern without interactions



(d) Curves of constant levels of the three-dimensional directivity pattern with interactions

Fig. 27 : Radiation of a system of four antennas on the spacecraft HERMES

COMPACT RANGES IN ANTENNA AND RCS MEASUREMENTS

by

B. AUDONE

Aeritalia - Gruppo Sistemi e Teleguidati
10072 Caselle (Torino) - Italy

Abstract

With the increased complexity and extended frequency range of operation model measurements and far field test ranges are no longer suitable to satisfy the demand of accurate testing. Moreover plane wave test conditions are required for Radar Cross Section (RCS) measurements which represent a key point in stealth technology. Compact Ranges represent the best test facilities available presently since they allow for indoor measurements under far field conditions in real time without any calculation effort. Several types of Compact Ranges are described and compared discussing their relevant advantages with regard to RCS and antenna measurements.

In parallel to measuring systems sophisticated computer models have been developed with such a high level of accuracy that it is questionable whether experiments give better results than theory. Tests performed on simple structures show the correlation between experimental results and theoretical ones derived on the basis of GTD computer codes.

1. Introduction

In both antenna and RCS measurements there is the need of generating uniform plane waves over large volumes. In case of real wavenumber k the plane wave is an equiphase surface perpendicular to \bar{k} ; plane wave uniformity additionally requires that the equiphase surface is also equiamplitude. The aim of RCS reduction consists in achieving low values of reflectivity and therefore RCS engineers are interested in low levels of return signals; viceversa antenna engineers want to achieve high levels of signals as an aim of their efforts directed to optimize antenna performances. RCS measurements are more critical not only because of this difference but also because the RCS test is a two way measurement while the antenna test is only one way; as a consequence in RCS measurement it is important to take more care of environment reflections which can seriously affect test results. Antenna test ranges do not generally require large frequency bandwidths because they are carried out at a discrete number of frequencies; viceversa RCS measurements require large bandwidths to achieve high resolution in ISAR (Inverse Synthetic Aperture Radar) techniques. This requirement creates several needs such as proper feed design and selection, anechoic chamber design, the verification of test performances over the entire frequency band of operation.

Antenna and RCS test ranges have been developed for many years and are in operation in many laboratories. They include:

- Outdoor ranges: they are used to perform measurements of large targets at distances which satisfy far field criteria. These facilities are easily established; their main disadvantages are lack of security, need of large ranges, total dependence upon weather conditions and ground reflections.
- Anechoic chambers: they are indoor facilities made by large rooms whose walls are covered with absorbing materials simulating free space conditions. The geometrical layout can be rectangular or tapered. Rectangular rooms were the first type of facility to be developed; tapered chambers have better performances at low frequencies.
- Near field test ranges: they are used for antenna measurements. The electromagnetic field transmitted by the antenna under test is measured in phase and amplitude by a probe antenna scanning a surface located at the distance of a few wavelengths. The surfaces may be planar, cylindrical or spherical. By means of well established theories which expand the electromagnetic field in planar, cylindrical and spherical elementary waves it is possible to compute the far field radiated by the test antenna transforming near field data into far field data.

- Compact Ranges (CR): these facilities, which will be discussed in this paper, use the collimating properties of one or more reflectors to produce a region of plane wave in which an antenna or the RCS of a target can be measured. Dielectric lenses can also be used to build CR facilities even if there are some practical limitations due to the required volume of homogeneous dielectric material and reflection interactions between lens surfaces; for these reasons they can be used for small test ranges.

2. Compact Ranges

There are three basic CR configurations:

- Offset reflector (Fig. 1a): there is only one reflector illuminated by the feed which is located outside the collimated beam.
- Dual offset reflector configuration: this type of CR uses the subreflector which can be located as in standard communication antenna system (Fig. 1b) or according to the front fed Cassegrain concept (Fig. 1c) which uses an offset hyperbola as subreflector and an offset parabola as main reflector.
- Dual cylindrical reflector configuration (Fig. 1d): it uses two parabolic cylindrical reflectors each of which collimates the electromagnetic wave in a different plane.

None of the proposed CR can produce a perfect uniform plane wave. The parameters to specify the quality of the quiet zone planarity are generally identified in amplitude taper and ripple, phase taper and ripple. The taper is determined by feed design; the ripple is caused by unwanted reflections within the CR anechoic room. Typical values are 5 to 20 electrical degrees of phase deviation, 0.25 to 1dB of amplitude taper and 0.1 to 1dB of peak to peak ripple. Other factors which define the CR performances are edge diffraction effects, feed pattern and polarization.

The CR performances are not only determined by the chosen reflector configuration but also depend on the overall system design which includes the type of chamber, the type and lay out of absorbing material, the measuring technique, the size of the unit under test, the feed. In the selection of the CR it is, therefore, important to adopt the overall integration approach in order to match the actual test requirements with the CR performances.

Compact Ranges have lower and upper frequency limitations. The lower limitation is due to the difficulty of achieving the required collimation performances and to edge diffraction effects.

There is much talk in the CR community about the edge design which represents one of the most controversial issue. Edge treatment is essential to reduce diffraction and to disperse energy from the edges into the absorbing material along the walls. The three main techniques adopted by the major manufacturers include the use of serrations, rolls (Fig. 2) and amplitude taper. The serrations are designed in such a manner that an amplitude taper is created towards reflector edges while keeping phase constant. Serrations provide a smooth transition from the high reflected field levels of the reflector surface to the zero field level immediately outside the reflector. This tapering action can be improved by coating the serration surface with resistive material. With the rolled edge the diffraction is reduced creating a phase taper toward the reflector boundary. The rolled edge must merge very accurately with the reflector structure: this can be achieved with a high quality control manufacturing process.

The main criticism to this solution derives from the consideration that the electromagnetic energy incident on the edges is redirected towards the walls of the anechoic chamber increasing the level of reflected energy in the quiet zone. The third method consists in the use of the combination of subreflector/reflector with the aim of shaping the amplitude and phase in order to provide a uniform illumination across the central part of the main reflector and to taper the energy towards the reflector boundary.

High frequency performances depend almost completely on the surface quality which is related to the manufacturing technique of the reflector. Experience has shown that small surface errors can cause unacceptable field variations. This limit becomes more stringent when the reflector is required to operate up to the frequency of 100 GHz; it has been derived that the maximum deviation of 0.007 wavelength can produce power level variations of 0.5dB in the quiet zone.

The CR configuration depends on the number and type of reflectors used. The single reflector CR (Fig. 1a) has a paraboloidal surface generated by rotating a parabola about its axis. The focus, where the feed is located, and the vertex lie on the axis. Feed blockage can be avoided by using offset geometry that is locating the quiet zone off of the paraboloid axis. When the feed is in the direct collimated beam of the reflector the feed receives back scattered spurious signals from itself and from the reflector; the latter one is particularly detrimental because very large. The feed shall be removed from the quiet zone in order to have as large the distance as possible to take the maximum advantage of reduced feed blockage. In this manner it is also possible to use more absorbing material behind the feed to reduce the direct feedback lobe radiation into the quiet zone. However increasing the feed offset increases the cross polarization unless the focal length is also increased.

A larger focal length can be achieved with a dual reflector system. This concept has several advantages:

- more uniform amplitude distribution due to larger focal length
- better figure of merit for aperture usage efficiency defined as the ratio between the total diameter of the reflector including the edges and the diameter of the quiet zone.
- lower cross polarization error.

Different types of dual reflector systems have been built. A manufacturer [6] uses the configuration subreflector/reflector to shape both amplitude and phase (Fig. 1b); the illumination is designed to be uniform across the central part of the main reflector and tapers towards the reflector edges. The feed is critical: its antenna pattern must remain constant across the entire frequency range of operation. Another critical area is the size of the subreflector which should be fairly large in diameter at low frequency (more than 12 wavelengths). Because of the required size of the subreflector it is quite possible that the main reflector is operated in near field conditions: this means that manufacturing tolerances of the reflector surfaces shall be tightly controlled because fabrication errors tend to give additive contributions.

A manufacturer [3] uses two parabolic cylindrical reflectors (Fig. 1d). The spherical wave from the feed is collimated in azimuth or elevation by the first cylinder and then collimated in elevation or azimuth by the second cylinder.

Because of the large focal lengths it is possible to achieve a high degree of uniformity in amplitude across the quiet zone. Similarly to what happens in far field test ranges the subreflector is seen within a cone of approximately 20 degrees which corresponds to a tolerance of 1.8dB. The copolar performance of this CR is asymmetric in the horizontal plane due to the differential space attenuation inherent to all offset geometry. In order to overcome this difficulty it has been proposed [9] to use an offset hyperbola as subreflector and an offset parabola as main reflector. The copolar pattern is symmetric in both the horizontal and vertical plane; the geometry induced cross polarization is zero for all polarizations.

3. Compact Range Evaluation and Use

Once the CR has been installed it is quite important to evaluate its performances measuring stray signals present in the quiet zone. This is generally done with one of the following methods:

- free space VSWR
- longitudinal pattern comparison

The former method is more suitable to detect stray signals from forward directions. Suppose that there are only a reflected signal with amplitude E_R arriving along the main antenna axis and a direct signal with amplitude E_D whose direction of arrival is shifted of angle α . The reflectivity level is calculated as

$$R = 20 \log \frac{E_R}{E_D}$$

When the test antenna is moved in and out of phase interference curve between E_D and E_R is recorded. The desired signal is detected at a pattern level L ; the total signal is the sum of

$$E_D 10^{L/20} \text{ and } E_D 10^{R/20}$$

The peak to peak variation σ is related to $I = R - L$ by the following equation

$$\sigma = 20 \log \frac{1 + \text{antilog}(I/20)}{\pm 1 + \text{antilog}(I/20)}$$

where the upper and lower signs in the denominator are chosen for I negative and positive respectively. In Fig. 3a the peak to peak variation σ (dB) is shown as a function of I (dB). Fig. 3b shows the peak to peak variation σ as a function of the pattern level L for reflectivity levels R equal to -20, -30, -40, -50dB. The antenna pattern comparison method more easily detects stray signals from transverse directions. The method which is based on the translation invariance of plane waves consists in recording the pattern of the receiving antenna at discrete points along the longitudinal line in the quiet zone. The recorded patterns are superimposed on one another so that the main lobe peak levels coincide. It is possible to evaluate the reflecting level R (as in the VSWR free space technique) by estimating the maximum variation at an aspect angle due to the in and out of phase interference technique between the direct signal and the reflected one.

The previous techniques more correctly applies to far field ranges where stray signal components can be considered to be of a plane wave character.

Within the CR reflected signals are generated in near field conditions by diffraction at the reflector edges, absorbers, feed and support structures. To cope with this situation it has been proposed [3,10,15] to adopt the spectral analysis method which additionally has the advantage of identifying the direction of the scatterers.

The electric field is measured with a probe antenna on a plane at $z = 0$ in the quiet zone in front of the CR main reflector. Suppose that there is an unwanted electric dipole source at $P(x_0, y_0, z_0)$ as shown in Fig. 4. The electric field on the probe antenna plane determined by this source is calculated according to the following equation

$$\bar{E}(\bar{r}) = - \int_{V'} \bar{\bar{Z}}(\bar{r}, \bar{r}') \cdot \bar{J}(\bar{r}') d\bar{r}' \quad (1)$$

where $\bar{\bar{Z}}(\bar{r}, \bar{r}')$ is the Green's dyadic function which is given by [20]

$$\bar{\bar{Z}}(\bar{r}, \bar{r}') = \frac{\omega\mu}{8\pi^2} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \frac{dk_x dk_y}{k_z} [\hat{u}_y(-k_z) \hat{u}_y(-k_z) + \hat{u}_z \hat{u}_z] e^{-j\bar{k} \cdot (\bar{r} - \bar{r}')} \quad (z < z') \quad (2)$$

where $\bar{k}^2 = k_x^2 \hat{u}_x + k_y^2 \hat{u}_y + k_z^2 \hat{u}_z$ $k = \omega \sqrt{\mu\epsilon}$

$$k_t = \sqrt{k_x^2 + k_y^2}; \quad \hat{u}_y(-k_z) = \frac{\hat{u}_t \times \bar{k}^-}{k_t}; \quad \hat{u}_z = \frac{\bar{k}^+ \times \hat{u}_t}{k_t}$$

In case $\bar{J}(\bar{r}') = \bar{J} \delta(x' - x_0, y' - y_0, z' - z_0)$ from (1) and (2) it follows

$$\bar{E}(\bar{r}) = \frac{\omega\mu}{8\pi^2} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \frac{dk_x dk_y}{(-k_z)} [\hat{u}_y(-k_z) \hat{u}_y(-k_z) + \hat{u}_z \hat{u}_z] \cdot \bar{J} e^{-j\bar{k} \cdot (\bar{r} - \bar{r}_0)} \quad (3)$$

where $\bar{r}_0 = x_0 \hat{u}_x + y_0 \hat{u}_y + z_0 \hat{u}_z$

By substituting

$$\bar{E}(k_x, k_y) = \frac{\omega\mu}{2} \frac{[\hat{u}_y(-k_z) \hat{u}_y(-k_z) + \hat{u}_z \hat{u}_z] \cdot \bar{J}}{(-k_z)} e^{+j\bar{k}^- \cdot \bar{r}_0} \quad (4)$$

In case $\bar{r} = x \hat{u}_x + y \hat{u}_y + 0 \hat{u}_z$ from (3) it follows

$$\bar{E}(x, y, 0) = \frac{1}{(2\pi)^2} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \bar{E}(k_x, k_y) e^{-j(k_x x + k_y y)} dk_x dk_y \quad (5)$$

$$\text{where } \bar{E}(k_x, k_y) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \bar{E}(x, y, 0) e^{+j(k_x x + k_y y)} dx dy \quad (6)$$

The integral (5) is calculated in the region $k_x^2 + k_y^2 < k^2$ where

$$k_x = k \sin \theta \cos \phi; \quad k_y = k \sin \theta \sin \phi; \quad k_z = \sqrt{k^2 - k_x^2 - k_y^2} \quad (7)$$

The propagation vector \bar{k} is at an angle θ with respect to the z axis and in a plane passing through the z axis that is at an angle ϕ with respect to the x axis. For $k_x^2 + k_y^2 > k^2$ the waves are evanescent and can be neglected if the probe antenna is at sufficient distance from the source.

The spectral analysis approach is rather meaningful because it identifies the direction of range scatterers in the C.R. Fig. 5 shows the plane wave spectrum of two sources located at $\theta = 45^\circ$ and at $\theta = 0^\circ$ and $\phi = 0^\circ$ with relative amplitudes of 0dB and -20dB respectively.

The vertical axis gives the amplitudes of the planewaves normalized to the peak of maximum value. The horizontal plane shows the direction of the sources: the circular lines represent increments of θ , the radial lines increments of ϕ . The amplitude of the electric field cannot be sampled over the whole probe plane but only in a limited region $-a/2 < x < a/2$ and $-b/2 < y < b/2$. Supposing to sample the x component of the electric field it follows

$$e_{tx}(x,y) = e_x(x,y) \cdot r(x,y) \quad (8)$$

where $e_{tx}(x,y)$ represents the field measured for a finite size aperture and

$$r(x,y) = \begin{cases} 1 & -a/2 < x < a/2, -b/2 < y < b/2 \\ 0 & \text{elsewhere} \end{cases}$$

In the spectral domain it follows

$$E_{tx}(k_x, k_y) = E_x(k_x, k_y) * R(k_x, k_y) \quad (9)$$

because the multiplication in the time domain corresponds to the convolution in the spectral domain according to Fourier transform theory. The evaluation of the Fourier integral

$$E_{tx}(k_x, k_y) = \int_{-a/2}^{a/2} \int_{-b/2}^{b/2} e_x(x,y) e^{j(k_x x + k_y y)} dx dy \quad (10)$$

is carried out by means of the Discrete Fourier Transform (DFT).

In the equivalent calculation of $E_{tx}(k_x, k_y)$ according to (9) the rectangle function $R(k_x, k_y)$ gives too high sidelobes which result in undesired peaks in the spectral domain. It may be advisable to select a more suitable window function which reduces the unwanted responses at the expense of reduced resolution. The role of window function becomes important in the case of point sources located at finite distances. From (4) by means of (7) considering only the x component of the electric field it follows:

$$E_x(k_x, k_y) = \frac{\omega \mu}{2} \frac{\hat{u}_x \cdot [\hat{u}_{//}(\cos \theta') \hat{u}_{//} \cos \theta' + \hat{u}_{\perp} \hat{u}_{\perp}]}{k \cos \theta'} e^{j k R_0 [\sin \theta \sin \theta_0 \cos(\phi - \phi_0) + \cos \theta' \cos \theta_0]} \quad (11)$$

where $\theta = \pi - \theta'$

The amplitude of $E_x(k_x, k_y)$ depends on the source orientation but is independent of the source distance; vice-versa the phase of $E_x(k_x, k_y)$ depends on the source distance R_0 . The plane wave spectrum obtained from samples measured in a finite aperture area is calculated by means of the convolution of (11) with the window function $R(k_x, k_y)$ as shown in (9). The convolution is an integral function; therefore when $k R_0$ is sufficiently large according to the saddle point approximation the major contribution is given by the direction $(\theta', \phi) \approx (\theta_0, \phi_0)$. The spectral response which has a peak in the direction of the point source has a shape determined by the window function. When the source approaches the aperture the spectral response broadens and loses definition. It has been calculated [15] that when R_0 is equal to 10 wavelengths the spectral response at relative amplitude of -10dB has a dispersion in the θ, ϕ plane of $\Delta\theta = 48^\circ$, and $\Delta\phi = 59^\circ$, while when R_0 is equal to 500 wavelengths $\Delta\theta = 10.5^\circ$ and $\Delta\phi = 8.3^\circ$.

Even in the best CR it is almost impossible to eliminate undesired signals due to scattering from the various parts of the anechoic chamber. This situation can be drastically improved by using time gated measurements which are based on the principle of differentiation of the signals due to the different times of arrival. This method which is applicable to RCS and antenna tests can be implemented by means of two techniques:

- grating CW systems
- instrumentation pulse radar

In the former case the transmitter is pulse modulated and the receiver is range (time) gated to measure the target signal. This technique uses the recently developed broadband, low reflection switches whose timing is selected to measure the wanted signal. The receiver is narrowband. This limits system performances according to two important aspects. The overall sensitivity is reduced by the square of the duty cycle. This problem can be partially solved by putting the receiver switch after the first low noise amplifier and before the receiver. In this way the input noise is attenuated by the switch of a factor equal to the duty cycle. Therefore the overall sensitivity is degraded only by the duty cycle [14]. The second limitation of the gated CW system is that the narrowband receiver slowly integrates the received energy throughout the entire pulse duration. In order to reach the steady state of the transmitted pulse the transmitted pulse train duration must be sufficiently long compared to the IF filter time constant. It has been suggested [16] that

$$2 \pi f_c NT \geq 3$$

where f_c is the 3dB cutoff frequency of the low pass filter, T is the PRI and N is the required number of pulses.

The instrumentation pulse radar uses a receiver with a wideband filter matched to the transmitted pulse. In order to obtain accurate and repeatable results [13] it is necessary to take into account many possible error sources which are related to pulse shape and repeatability. Receiver performances depend on thermal noise, signal compression, phase (or I/Q) measurement accuracy, digitization error, receiver bandwidth. Transmitter performances depend on waveform purity, turnoff time and output level versus frequency.

In RCS measurements it is required to resolve target returns in range. This can be done reducing the pulse duration of the waveform which, as a consequence, reduces radar detection range. To overcome this difficulty it is necessary to modulate the waveform to increase the bandwidth suitable for the wanted range resolution. Processing the received waveform is, then, utilized to compress the radar waveform into a smaller range interval; this operation creates range sidelobes which can reduce the measuring system dynamic range because of the obscuration effect of small targets near larger ones.

In antenna measurements the pulse duration T (expressed as a length) is established on the basis of the following rule (Fig. 6).

$$2D < T < |L_U - L_W|$$

where D is the maximum size of the antenna under test, L_U is the path length of the unwanted signal and L_W is the path length of the wanted signal. It is clear that when the L_U does not differ much from L_W the wanted signal cannot be discriminated unless the pulsewidth is sufficiently narrow to meet the previous criterion. Time gating reduces stray signals due to feed backlobe radiation and some anechoic chamber reflections in the same manner as absorbing material does. It cannot remove reflector edge or feed support diffraction effects from the collimated signal.

In the evaluation and use of the CR it is important to characterize its performance in terms of the efficiency η with which power is delivered from the input terminal of the feed to the quiet zone. The power density P_D in the quiet zone is given by

$$P_D = \eta \frac{P_T}{A} \quad (12)$$

where P_T is the power at the feed terminal inputs and A is the quiet zone area. The power reflected by the target with RCS σ is

$$P'_R = P_D \sigma = \frac{P_T \sigma \eta}{A} \quad (13)$$

The reflected power is reradiated isotropically

$$P_{DR} = \frac{P'_R}{4\pi} = \frac{P_T \sigma \eta}{4\pi A} \left(\frac{W}{\text{ster}} \right) \quad (14)$$

The received power is given by

$$P_R = P_{DR} \cdot \Delta\Omega \quad (15)$$

where $\Delta\Omega$ is the solid angle passband of the CR. It has been estimated [6] that $\Delta\Omega$ is approximately given by the area within the 3dB beamwidth of the far field pattern of the antenna or

$$\Delta\Omega = \left(\frac{\lambda}{D_E} \right)^2 \quad (16)$$

It follows

$$P_R = \frac{P_T \eta \sigma \lambda^2}{4 \pi A D_E^2} \quad (17)$$

where D_E is the effective diameter of the reflector. In case of dual shaped reflector system assuming well designed feeds it is possible to achieve efficiencies greater than 90%. In case of front fed systems typical values of efficiency are in the range of 35 to 40% [19].

4. Experimental verification of simple theoretical cases

A series of measurements have been carried out in the Compact Range with two main objectives:

- to verify the overall performances of the test range
- to validate a UTD computer code with simple canonical cases

A UTD computer code to determine the patterns of antennas located on aircraft structure has been developed. The free space radiation pattern of a source is degraded to some extent when the antenna is mounted on the aircraft since energy is reflected and/or diffracted mainly by wings, fuselage, tail and also by other assemblies such as the undercarriage, flaps, slats, external stores, pylons and other antennas. The objective of antenna location studies consists in finding out the best position where the deterioration of the antenna radiation pattern is minimized mainly in that part of the sphere where the probability of receiving or transmitting is high. The UTD approach is ideal for this case: ray optical techniques are used to determine components of the field incident on or diffracted by the various structures.

A UTD computer code to determine the patterns of antennas located on aircraft structure has been developed. The total field in a given direction is obtained by taking the sum of the field resulting from a number of different scattering mechanisms. Each component is determined by tracing the ray through the appropriate geometrical path and by using the UTD to compute the magnitude and phase of the field if it has not been shadowed. The code computes up to a second order field contribution, that is, it can include fields for ray paths that interact with at most two scattering centres.

Complicated structures can be simulated by arbitrarily oriented perfectly conducting or dielectric plates, an infinite ground plane, and perfectly conducting elliptic cylinders. The field computation is divided into five sections. The first section computes the incident field, that is the direct field from the source to the observation point. The second section contains the major scattered fields associated with the individual flat plates and the interaction between the different plates. These include the single reflected fields, the single diffracted fields, the reflected-diffracted fields, and the diffracted-reflected fields.

The third section analyzes the major scattered fields associated with the finite elliptic cylinder. This includes the uniformly reflected fields, the uniformly diffracted fields, the reflected fields from the endcap rims, the fields uniformly reflected and/or diffracted and then uniformly reflected and diffracted from another cylinder. Presently, the second order cylinder interactions can only be used for parallel cylinders.

The fourth section contains the major scattered fields associated with the interactions between the plates and cylinders. This includes, at present, the fields reflected from the plates then scattered from the cylinder and the fields reflected or diffracted from the cylinder then reflected from the plates. The creeping wave masking is present. The fifth section contains the double diffracted fields associated with the interaction among plates.

The field pattern of an antenna in presence of scatterers can be obtained by scanning with fractional step angle, to increase the definition, when this operation is physically significant.

The following cases have been studied by comparing experimental and theoretical results:

- monopole and an edge
- monopole and a double edge
- monopole and a smooth step
- monopole and a double step
- monopole and a cylinder
- monopole, a cylinder and a flat plate installed on it in various locations.

The tests have been carried out in a CR installed in an anechoic chamber (12 m long, 8 m wide and 7 m high). The receiving-generation system is the SA 2020 Antenna Analyzer. The CR has an offset paraboloidal reflector. In Fig. 7 the source which is a short monopole parallel to the edge is located at the distance of one wavelength from the nearest point of the edge. The pattern of E_θ is measured along ϕ . The computed pattern takes into account the double diffraction interaction between edges. It is interesting to notice the very good agreement after 180° .

Fig. 8 shows the comparison between the computed patterns with and without the double diffracted field.

In Fig. 9 the scattering between the field source and the double step is shown. The polarization of the electric field is perpendicular to the horizontal scatterer edges. The pattern is measured in the xz plane. The distance between the source and the nearest plate is 1.8 wavelength. At $\theta = 0$ the field is low because of the destructive interference of the direct and reflected rays. The coincidence of calculated and measured results at $\theta = 45^\circ$ is very good.

Fig. 10 shows the comparison between the computed patterns with and without the double diffracted field. One can observe how the discontinuity at the shadow boundary is removed when the double diffracted field is taken into account.

Fig. 11 shows the interaction between the source and a smooth step. In this case the diffracted field coming from the first edge to the second one is in grazing incidence condition. The large difference at about 180° is due to the lack of the triple diffracted field from the third edge.

Fig. 12 shows the comparison between the cases with and without the double diffraction.

Fig. 13 shows the pattern of E_θ of a short monopole in presence of a structure made by a cylinder and a flat plate.

In Fig. 14 the comparison between the cases with and without plate cylinder interaction is shown.

Fig. 15 shows the pattern of E_θ due to an array of two monopoles. The agreement is very good in the angular region between 0 and 180° where the contributions reflected from the plate and then scattered from the cylinder and viceversa are important.

5. Conclusion

The CR gives almost ideal test conditions with the advantage of the wide frequency range of operation in an anechoic chamber of limited size. The CR concept has been interpreted in several manners with solutions affected by advantages or disadvantages mainly in relationship to the specific measurements the user intends to carry out. Tests on some simple canonical articles have been performed to verify the validity of the test range and the computer code.

References

1. Compact ranges get radar cross section data abundance. John Pustai, Microwaves and RF, May 1987.
2. Compact range technology and the marketplace. R.C. Johnson and C.E. Ryan. Microwaves and RF, May 1987.
3. Troubleshooting limitations in indoor RCS measurements. Berth Schluper and J. Vokurka, Microwave and RF, May 1987.
4. Dual chamber design reduces quiet zone ripple/taper errors. W. Burnside. Microwave and RF, May 1987.
5. Prime focus feed key to operation of largest range. John Pustai. Microwave and RF, May 1987.
6. A new approach to radar Cross section Compact Range. Terry Harrison. Microwave Journal, June 1986.
7. A serrated edge virtual vertex compact range reflector. Doren W. Hess and Kevin Miller. Proceedings of the 11th Estec Antenna Workshop on Antenna Measurements. June 1988.
8. A millimetre wave compact antenna range. P. Mac Nair, A.D. Oliver and C.G. Parini. Proceedings of the 11th Estec Antenna Workshop on Antenna Measurements. June 1988.
9. Development of an optimized Compact Test Range. E. Dudok, D. Fasold and H.J. Steiner. Proceedings of the 11th Estec Antenna Workshop on Antenna Measurement. June 1988.
10. Evaluation of antenna test ranges by using plane wave spectral analysis. P.A. Beckman. Proceedings of the 11th Estec Antenna Workshop on Antenna Measurements. June 1988.
11. Radar Cross Section Measurements. Robert B. Dybdal. Proceedings of the IEEE, April 1987.
12. The Ohio State University Compact Radar Cross Section Measurement Range. E.K. Walton and J.D. Young IEEE Transactions on Antenna and Propagation, November 1984.

13. Radar cross section and imaging measurement systems. William G. Swarner and Richard H. Bryan MSN, April 1988.
14. A new perspective on radar cross section measurement system. B. Audone, A. De Logn and F. Furini. Proceedings of the 11th Estec Antenna Workshop on Antenna Measurements.
15. A planewave spectral range probe. R.D. Coblin. 10th Annual Antenna Measurement Techniques Association Proceedings September 1988.
16. Performance of gated RCS and Antenna measurement L.R. Burgess and D.J. Markman. 10th Annual Antenna Measurement Techniques Association Proceedings, September 1988.
17. Time gating of antenna measurement. D. Hess and V. Farr. 10th Annual Antenna Measurement Techniques Association Proceedings, September 1988.
18. Analysis of Compact Range reflectors with serrated edges. K. Miller and R.W. Krentel. 10th Annual Antenna Measurement Techniques Association Proceedings, September 1988.
19. Transfer efficiency of the Compact range. R.W. Krentel. 10th Annual Antenna Measurement Techniques Association Proceedings. September 1988.
20. Radiation and scattering of waves. L. Felsen and W. Marcuvitz. Prentice Hall, 1973.

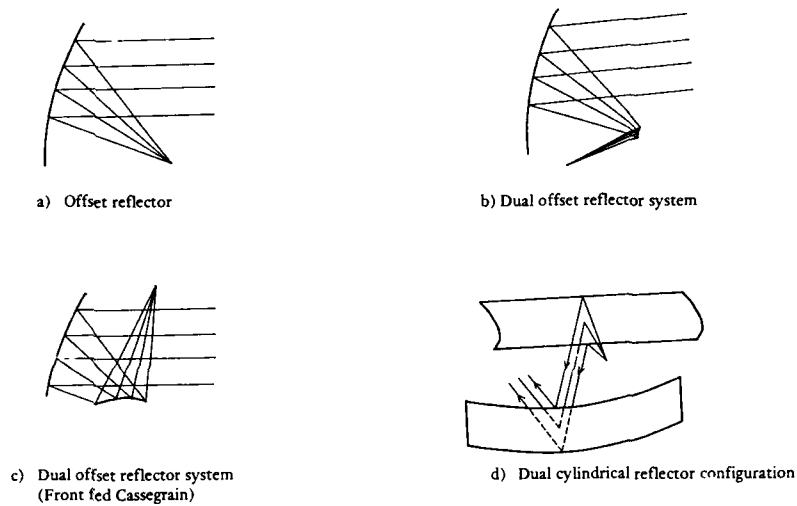
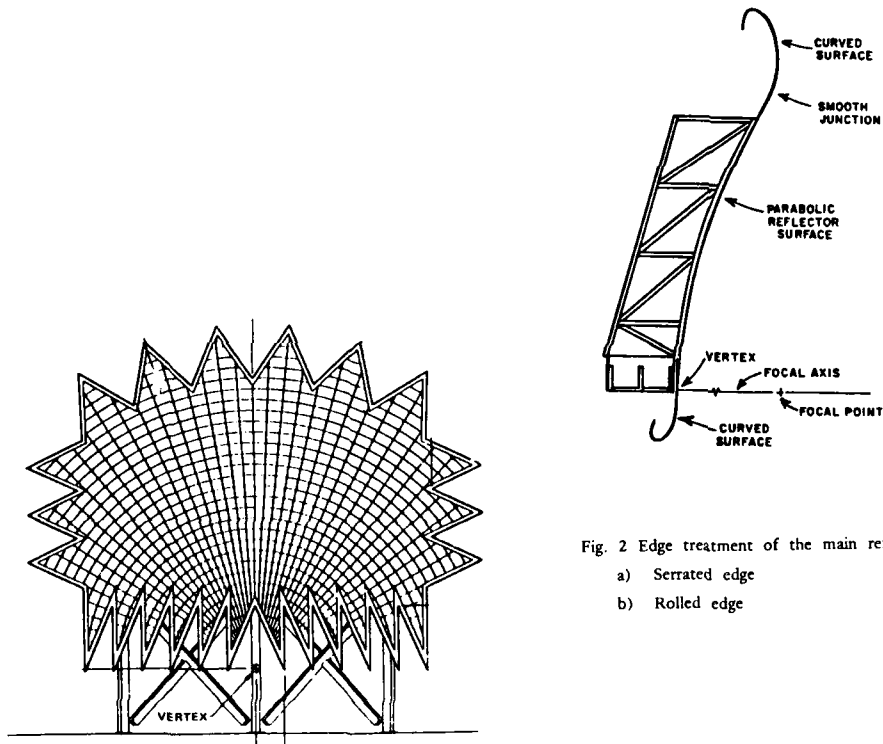


Fig. 1 Basic Compact Range configurations



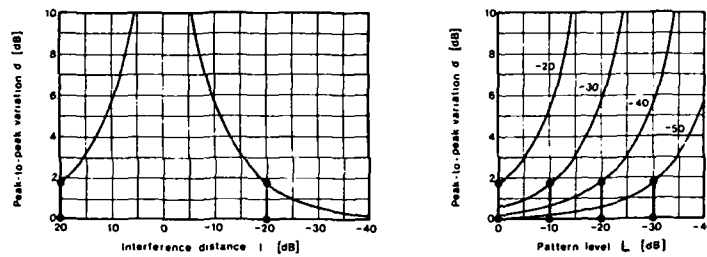


Fig. 3 Evaluation of stray signal levels
 a) peak to peak variation σ versus I
 b) peak to peak variation σ versus L

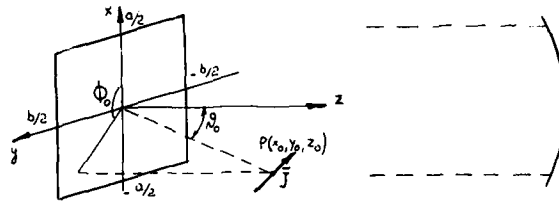


Fig. 4 Unwanted electric dipole source in front of the reflector

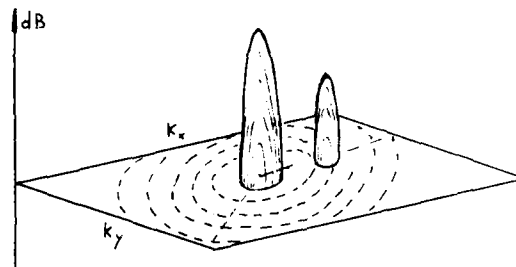


Fig. 5 Plane wave spectral representation

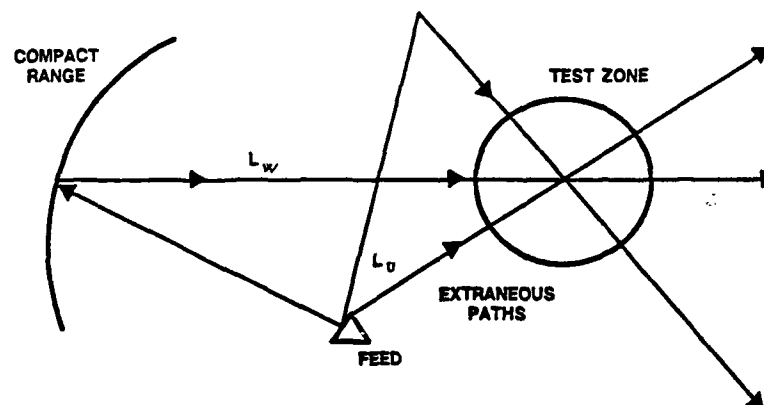


Fig. 6 Wanted and unwanted signal paths

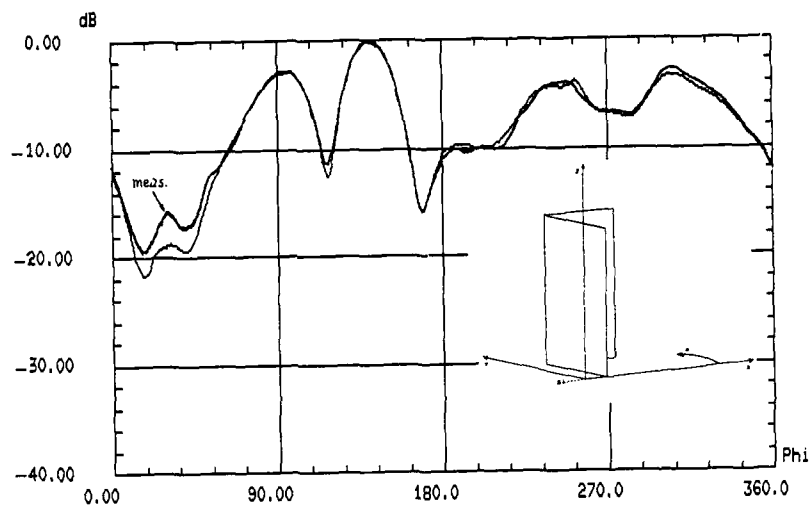
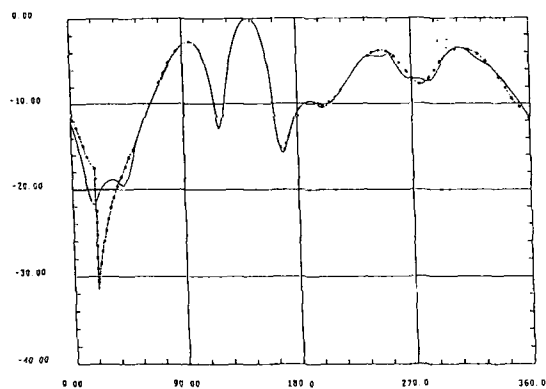


Fig. 7 Pattern of a short monopole parallel to the edge



WITH (——) AND WITHOUT (-----) DOUBLE DIFFRACTION INTERACTIONS

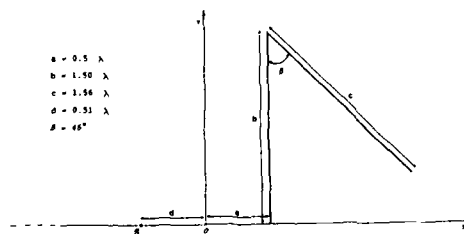


Fig. 8 Comparison between the computed patterns with and without the double diffracted field (monopole/edge)

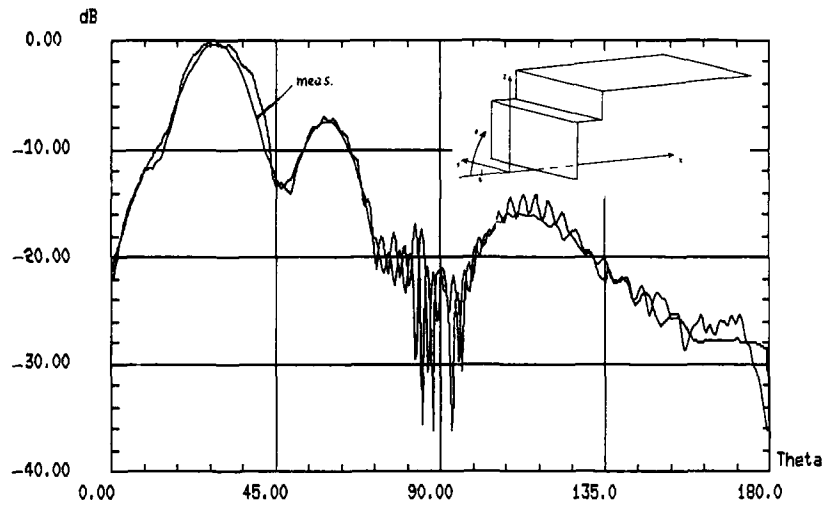


Fig. 9 Pattern of a short monopole in presence of a double step

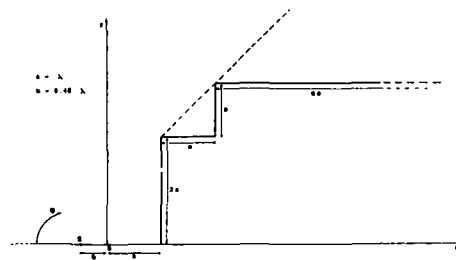
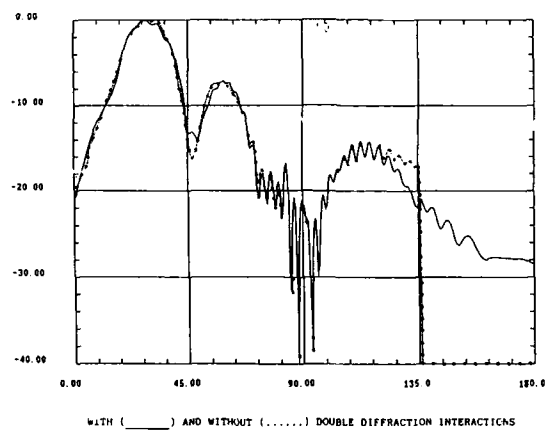


Fig. 10 Comparison between the computed patterns with and without the double diffracted field (monopole/double step)

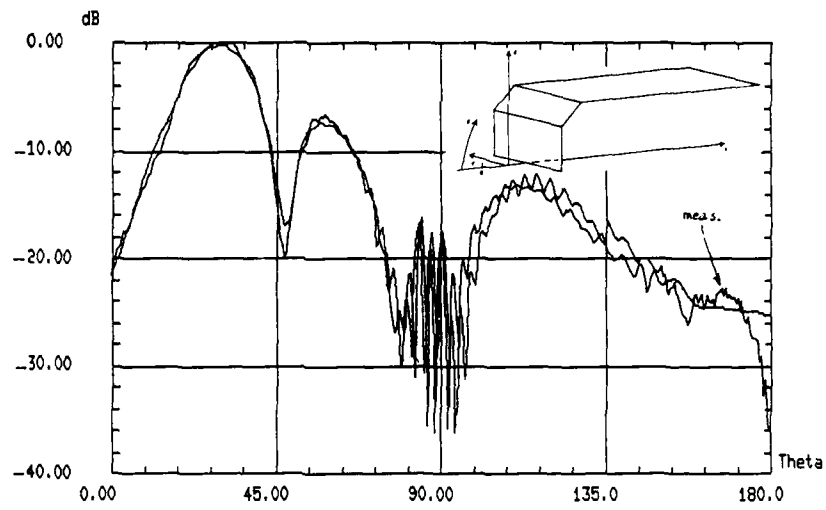
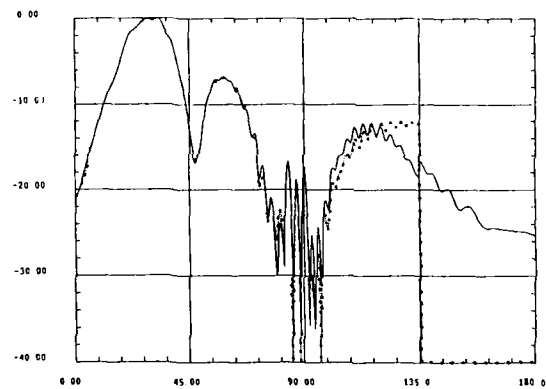


Fig. 11 Pattern of short monopole in presence of a smooth step



WITH (—) AND WITHOUT (---) DOUBLE DIFFRACTION INTERACTIONS

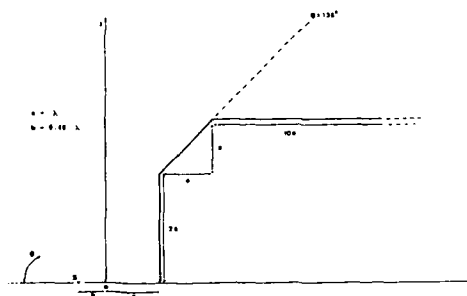


Fig. 12 Comparison between the computed patterns with and without the double diffracted field (monopole/smooth step)

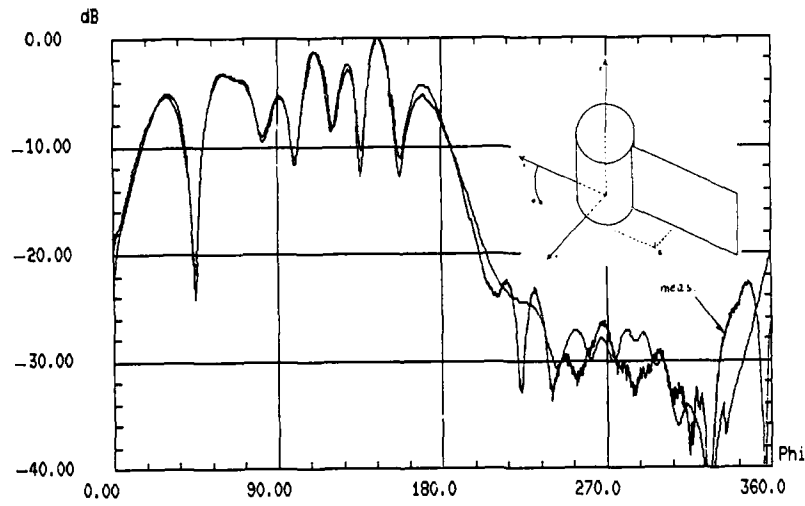


Fig. 13 Pattern of a short monopole in presence of a cylinder and a flat plate

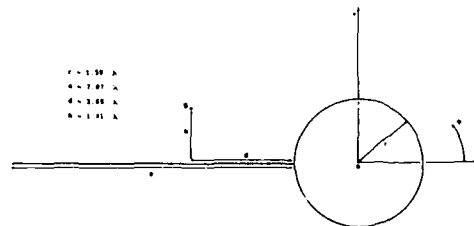
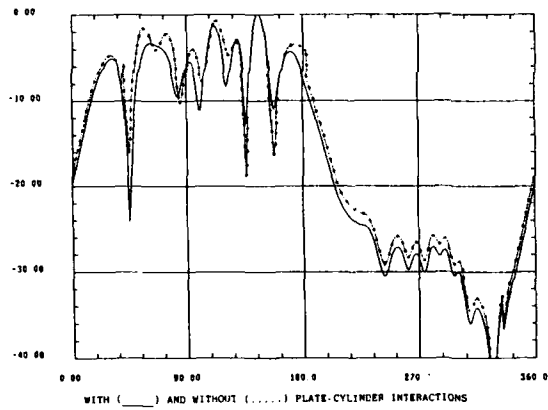


Fig. 14 Comparison between the computed patterns with and without the double diffracted field (monopole/cylinder and plate)

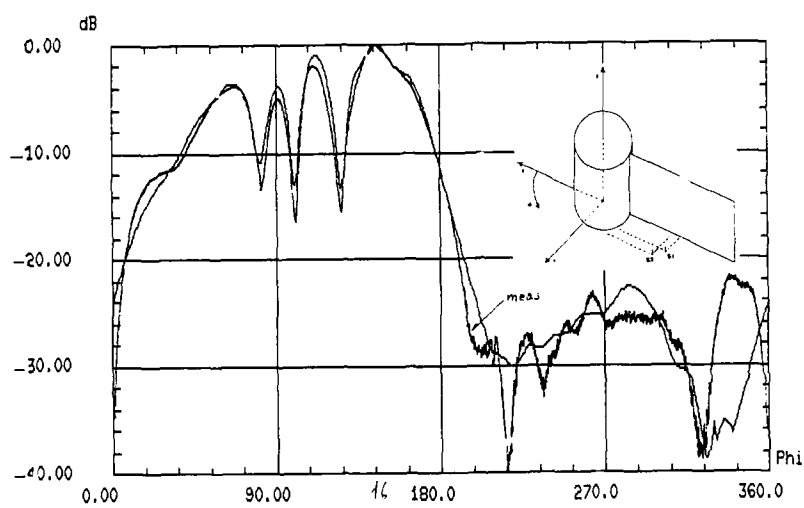


Fig. 15 Pattern of an array of two monopoles in presence of a cylinder and a flat plate

ANTENNA INSTALLATION ON AIRCRAFT:

THEORY AND PRACTICE

Author:
T.J.MURPHY
MANAGER,
MICROWAVE ANTENNAS.

BRITISH AEROSPACE (DYNAMICS) LTD
FPC 077, P.O.BOX 5,
FILTON HOUSE
BRISTOL BS12 7QW.

SUMMARY

This presentation reviews the problems to be faced in the integration of an antenna system onto an aircraft. In the discussion the techniques, both mathematical and empirical, used to assess radiation pattern performance are placed in the context of the requirement on the airframe as a system. A systematic approach to the integration task is described which experience has shown to be suited to the overall task.

1. INTRODUCTION

The task of installing an antenna system on an aircraft is one which requires a wide range of skills. Other presentations in this Lecture Series have introduced a number of techniques for predicting the performance of antenna systems and some of these will be mentioned in passing here. The main objective of this discussion is to illustrate the range of skills and expertise needed in order to successfully integrate an antenna system into an airframe.

Almost any modern aircraft, be it civil or military, will carry a significant number of electronic (strictly electro-magnetic radiating) systems. These serve purposes ranging from simple voice communications to multi-function radars. Each of these systems will represent, to the operator, an increase in the price of using the airframe both in simple capital outlay and more importantly in increased maintenance and fuel costs and reduced carrying capacity. Since even the simplest system threatens significant costs of ownership it is essential that performance be optimised.

Part of the optimisation process will be to ensure that the antennas are placed in appropriate locations. However it must be remembered that on small airframes, in particular, it may be that the radiation pattern shape is not the only driver.

As an example of the number of antennas used on a modern airframe the BAe 125 commuter jet is shown figure 1. This carries as standard some 16 antennas (these are listed in table 1.1) to provide the communications and navigation facilities needed to use today's crowded airways. Clearly the additional mission related systems on military aircraft can only add to the problems of installation and operation. For example a naval ASW helicopter might have up to 40 antennas.

It is often the case that the multitude of antenna systems cannot all be mounted in the most desirable position for the particular function. The need is rather that the airframe system as a whole should be optimised. It is this view that will be developed in the following.

Table 1.1: BAe-125 ANTENNA SYSTEMS

ANTENNA SYSTEM	FREQUENCY (in MHz)	FUNCTION
GLIDESLOPE (2 OFF)	328 - 336 MHz	Vertical component of ILS
VOR/LOCALISER (2 OFF)	108 - 118 MHz	- Extraction of beacon bearing - horizontal component of ILS
RADAR	I - J Band	Weather warning
MARKER	75 MHz	Distance marks for ILS
ATC (2 OFF)	1 - 1.1 GHz	Secondary surveillance Radar
RAD.ALT	C Band	Height measurement
DME (2 OFF)	960 - 1260 MHz	Distance and range to beacon
VHF COMMS (2 OFF)	118 -150 MHz	General communications
HF COMMS	2 -30 MHz	Long range communications
ADF (2 OFF)	200 - 1600 KHz	Airborne direction finding

2. CONSTRAINTS

The design and installation of aircraft antennas are subject to some very basic, and at times inconvenient, constraints. These will limit the options available both to the antenna designer and the airframe manufacturer. A few of these are listed here:-

- (a) The aircraft must remain able to fly safely. This means that the antenna's weight and location must not compromise airframe trim limits; the size should not impose excessive drag or structural loads and it should not so disturb the airflow over control surfaces as to effect stability.
- (b) The antenna installation should not compromise the aircraft role, for example by reducing flight duration or hindering the carriage or deployment of stores.
- (c) The antenna system should not compromise the performance of other systems. For example the performance of the ILS (Instrument Landing System) is critical to flight safety and the addition of a poorly designed new installation could markedly degrade its performance.
- (d) The antenna site must be sufficiently close to the associated avionics for the losses and weight of the necessary feeder cable to be acceptable.

It is clear that, in a military environment, the rigour attached to some of these constraints would be less than in the civil area. Examples of this can be seen in the antenna installations on the AWACS E3A and the ill-fated NIMROD AEW which were judged to be sufficiently important to justify the radar antennas dominating the airframe structure. However it is important that, whatever the compromise in performance, the operator be given the information which will allow the necessary decisions to be taken in a balanced manner.

3. INSTALLATION

In this part of the discussion the tacit assumption has been made that the airframe is new and as a consequence the interaction with other antennas can be ignored. Some sites will tend to have been assigned automatically by the aircraft designer, for example the radar usually goes in the nose (behind an aerodynamic cover) of a civil aircraft.

3.1 GETTING STARTED

The first step in any aircraft installation exercise is always data gathering. It is essential that as much detail as possible is gathered on the airframe structure, flight profile (speed, height and attitude limits), system electrical and mechanical requirements and any other, special considerations. The structure of the airframe is particularly important since the loads generated by the antenna during flight will have an effect on stressing levels and must be catered for. In addition, if the aircraft is pressurised the antenna installation must actively provide for maintaining the integrity of the seal.

The aircraft flight profile will define both the mechanical strength of the antenna and its mounting and the coverage demanded. In addition a highly manoeuvrable aircraft may have a requirement that the antenna size be limited to preserve rates of turn.

Coupled with manoeuvre is the question of areas of coverage. A civil aircraft will not normally exhibit large excursions from level flight, say up to 30° in roll and pitch, for significant portions of its flight. On the other hand a small military aircraft could be expected to experience roll angles in excess of 65° for significant periods of time. If the antenna system is expected to operate in this attitude, and is linearly polarised, then signal reductions of more than 6dB can be expected. Other effects such as screening due to wings will also have to be considered during manoeuvre.

The last paragraph introduces the question of electrical performance. It is necessary to identify the role of the system being considered to ensure that appropriate margins are provided or the consequences of failing to provide these margins.

The type of siting constraint which might come under the heading "special" are best shown by example.

On one civil airliner there was a requirement that all antennas on the underside were flush fitting so that ground crew did not hit their heads on them.

On another civil airframe there was a requirement that all the underneath antennas on the fuselage had 100mm long stand-offs for the connectors. This was to keep the connections out of the bilge water which accumulated during flight.

3.2 DEFINE ANTENNA COVERAGE REQUIREMENT

During this phase, the information on the airframe and the system are brought together to define the sectors about the aircraft where antenna coverage is needed, and the levels of performance required in each region.

3.3 SELECT ANTENNA TYPE

The selection of an antenna type will only be possible when the required coverage is defined. With a few notable exceptions the monopole has been the element of choice for aircraft systems. However there are at least 16 identified different forms of monopole available to meet trade-offs in size, bandwidth and mechanical robustness. There are also a number of antennas with similar mechanical shapes which offer different radiation pattern structures, for example cardioid rather than omnidirectional. It is in this area that special constraints such as those mentioned in 3.1 will come into play.

3.4 ANTENNA INSTALLATION PLAN

In this phase an antenna installation plan is prepared, in conjunction with the airframe design authority. This must include provision for the effects of antenna loads on the airframe, location of associated avionics, cable losses and routing and system priority in terms of aircraft handling and role. These will be discussed in greater depth later.

In this activity it is necessary to have a method of approximating the antenna coverage to ensure that the sites initially selected are suitable to meet the requirement. It will usually be necessary at this stage to retain a number of possible configurations to reflect different compromise solutions.

A simple modelling tool is useful in this area to sort the good from the bad quickly. The principle need here is for speed and ease of use; the achievement of mathematical rigour and precision is not really necessary.

3.5 PERFORMANCE PREDICTION

The final choice as to the preferred installation option must be confirmed by some firmly based method of assessment.

The time honoured method is scale model measurement (used at least since the 1940's) where a scale model of the airframe is fitted with a suitable scaled antenna in the intended location and the radiation patterns of the whole measured.

In recent times mathematical techniques have come to the fore and, as computer power becomes cheaper, will gain in importance. It must be noted however that there remain some limitations as to the size and complexity of the systems

which these techniques can handle. In particular presently available techniques find it difficult to handle airframe dimensions of a few wavelengths. These often occur on a typical airframe for frequencies in the 100 MHz to 1 GHz range.

3.6 ANTENNA ENGINEERING

In this phase the actual design of antenna will be settled upon. The first step should always be a search of the marketplace for an existing design. It is admittedly necessary on occasion to design a new element for a particular application but an existing design will often give cost savings.

3.7 QUALIFICATION

The antennas chosen, or developed, must be proven to be mechanically adequate for the proposed installation and to perform electrically, within acceptable tolerances, in the expected environment.

3.8 TRIALS

The antenna installation must be proven by trials on an actual aircraft. This should be part of a full system trial in which the whole system is exercised to prove that all aspects of the installation function and that there are no unexpected interactions with other aircraft equipment.

4. VALIDATION OF THE ANTENNA INSTALLATION

At each step in the installation of an antenna system it is necessary to show evidence that the sites chosen are likely to be acceptable. In this section the principal alternatives for gaining confidence in the antenna radiation patterns are discussed. A total of five methods have been included:-

- Measurements taken on the ground with the antenna fitted to an airframe or full size model.
- Measurements taken in flight with the antenna fitted on the target airframe.
- Measurements on a (reduced) scale model airframe.
- Modelling using mathematical techniques such as G.T.D. (Geometrical Theory of Diffraction) or wire-grid (i.e. N.E.C.) which offer a degree of rigour and precision.
- Predictions based on experience such as the technique suggested by Cary which uses patterns for monopoles mounted on groundplanes of standard size to synthesise the installed radiation pattern.

In any discussion of the various approaches it is necessary to recognise the errors inherent in each and to make choices appropriate to the stage in the development programme.

4.1 FULL SCALE GROUND MEASUREMENTS

Ground based measurements fall basically into two categories, moderately expensive or very expensive. The amount of information which can be gleaned reflects the price difference.

In the first approach the airframe stands on, or close to, the ground. Pattern data can only be recorded as azimuth patterns and the range of elevation angles which can be examined is strictly limited. Strictly if antennas on the underside are to be examined then the airframe, or model, must be turned upside down.

In the second approach a specially prepared airframe or model is mounted on a positioner system which allows the full sphere of coverage to be examined. This technique has been used in the USA and France for relatively small aircraft.

In any ground based system the major area of error is likely to be caused by multipath, reflections from the ground and surrounding structures. This means that the range site and geometry must be carefully chosen and it may need the use of diffraction fences to reduce errors to an acceptable level. A detailed probing of the amplitude variation across the aircraft will probably be required to demonstrate a satisfactory range design. It must also be remembered that for some systems significant ranges will be needed in order to enable far-field performance to be established.

Table 4.1 summarises typical errors for this type of measurement.

Table 4.1: Errors in Full-scale ground based measurements

Errors for three antenna gain levels are included.

Source	Level		
	5.dBi	-5 dBi	-15.dBi
Reference antenna gain	± 0.2	± 0.2	± 0.2
Diffraction fences	± 0.5	± 1.5	$+4.6/-7.0$
Cable and connectors	± 0.1	± 0.1	± 0.1
Dynamic Range	± 0.1	± 0.15	± 0.2
Plotter/Digitiser	± 0.1	± 0.1	± 0.1
Worst case absolute error	± 1.0	± 2.05	$-7.6/+4.6$
Mean error in absolute gain	± 0.4	± 0.9	N.A
Worst case relative error	± 0.8	± 1.85	$-7.5/+4.6$
Mean error in relative gain	± 0.3	± 0.9	N.A

4.2 MEASUREMENTS IN FLIGHT

Flight trials to assess radiation patterns on aircraft (excluding helicopters) are normally carried out using one of two flight profiles; orbits or daisy-patterns. It is a characteristic of these that the pilot has a very significant workload imposed on him if the results are to be worthwhile.

Both types of flight profile involve centring the aircraft track on a convenient reference point at a moderate range from the ground station. The choice of this point and the track to be flown will be influenced by the terrain in the region.

When flying orbits the aircraft follows a nominally circular track, at a fixed bank angle, whilst transmitting continuously. It is inevitable that range corrections will have to be made for this type of measurement since the orbit radius will be fixed by the bank angle and speed; small bank angles and/or high speed will lead to a large radius of turn.

In measurements using the daisy-pattern track the aircraft is flown in a series of straight lines atop the reference point and transmits briefly at the crossing point. A complete 360° azimuth plot at 10° intervals requires a total of 37 crossings of the reference point; the 37th serves to give a check that equipment has not drifted excessively during the flight.

The principal causes of error divide into two classes; the first is pilot related and the second system related. The pilot's workload is very heavy with extreme precision demanded if positioning errors are to be kept small. System errors include items such as range corrections, multipath, heading (in pitch, roll and yaw) and equipment drift with time.

A summary of the errors to be expected is very difficult to produce but experience has shown that in the best conditions angular errors of the order of $\pm 2^\circ$ and amplitude uncertainties of $\pm 3\text{dB}$ can be expected for straight and level flight. The errors for regimes involving high bank angles will grossly exceed these levels.

4.3 REDUCED SCALE MEASUREMENTS

Scale model measurements rely on the fact that if the dimensions of a structure are reduced by a given factor the radiation characteristics of the system will be preserved if the operating frequency is raised by the same factor.

The attraction of this is two-fold; the principal reason is the savings in cost which can accrue and the second is its convenience, including where necessary the possibility of working in a secure environment.

In general the results obtained will be as good as the precision of the model will allow. The major areas of difficulty experienced with this method are in the modelling of dielectric structures such as radomes and the physical problems of achieving an accurate antenna model at large scale factors.

Table 4.2 summarises the likely error sources and their magnitudes for a monopole type antenna on an airframe at a scale factor of the order of 6, i.e. a 500 MHz antenna scaled to 3 GHz.

Table 4.2: Errors in Scale model measurements

Errors for three antenna gain levels are included.

Source	Level		
	5.dBi	-5.dBi	-15.dBi
Reference antenna	± 0.2	± 0.2	± 0.2
Full size antenna directivity	± 0.1	± 0.1	± 0.1
Site/chamber reflectivity	± 0.1	± 0.3	± 0.8
Cable/connectors (at scale)	± 0.2	± 0.2	± 0.2
Dynamic range	± 0.1	± 0.1	± 0.1
Plotter/digitisation	± 0.1	± 0.1	± 0.1
Worst case absolute error	± 0.8	± 1.0	± 1.5
Mean error in absolute gain	± 0.3	± 0.35	± 0.55
Worst case relative error	± 0.5	± 0.7	± 1.2
Mean error in relative gain	± 0.25	± 0.3	± 0.5

4.4 MATHEMATICAL MODELLING

A number of techniques exist for the mathematical modelling of the performance (radiation pattern) of antennas on an airframe. The techniques discussed elsewhere in this Lecture Series are clearly of relevance in this regard. It must be stated that again it is a case of the validity of the results reflecting the time and money dedicated to their production as long as an appropriate technique was used.

It is of particular importance that the scale of the structure in wavelengths be carefully considered in the choice of technique. There is a further concern in the modelling of dielectric structures.

In the summary at the end of this chapter there are some error estimates for a particular implementation of the GTD on a small airframe at a nominal frequency of 1 GHz.

4.5 SIMPLIFIED MATHEMATICAL MODELS

Some 10 years ago the advent of desk-top computers made it possible to consider simple models where accuracy was not the prime concern. This type of model is particularly useful during the early stages of a siting study where it is necessary to rapidly identify possible sites and their limitations.

There are a large number of possible implementations of this approach and the one included in the summary has been based on a method proposed by Cary in the 1950's.

It must be stressed that this type of approach is not advocated as an accurate method for the prediction of radiation patterns but rather as a qualitative indicator for good or bad positions. This lack of precision is reflected in the fact that results can be usefully be obtained working from a scale drawing of the airframe and ignoring almost all detail.

4.6 SUMMARY

There is never a clear cut answer to the question "Which is the best approach". The answer must always be coloured by considerations of cost, required precision, frequency of operation, size of airframe and the importance of system interactions.

The summary table, Table 4.3, given here has been cast for a nominal frequency of 1 GHz with a medium to small airframe. The result of the exercise was that scale modelling would probably be preferred.

It must be restated however that at the end of the day full scale ground or airborne trials will be needed to clear most installations.

The place of mathematical tools is strengthening as computation costs moderate and it is likely that in the near future at least some of the advantage of scale work will disappear. This is already the case for frequencies significantly higher than about 3 GHz since the difficulties of constructing scale antennas rapidly become intractable.

Table 4.3: Summary of estimated errors

	Full scale (Ground)	Full scale (Flight)	Scale Model	Mathematical Model (GTD)	Cary Method
FACTORS					
Surface accuracy	N.A	N.A	1/10	1/10	1/2
Dielectrics included	YES	YES	NO	NO	NO
Dimensional accuracy	N.A	N.A	±2%	±2%	±2%
360° Azimuth coverage	YES	YES	YES	YES	YES
Elevation range	near 0°	near 0°	±90°	±90°	±90°
Time for 1 azimuth pattern	2-3 min.	30-90 min.	2-3 min.	-	-
Cost	High	Very High	Medium	Medium	Low
ERRORS (dB)					
Mean (Absolute Gain)					
- at 5dBi	±0.4 (±1.0)	not	±0.3 (±0.8)	-	Envelope
- at -5dBi	±0.9 (±2.5)	defined	±0.35 (±1.0)	-	predicted
- at -15dBi	several dB		±0.55 (±1.5)	-	to ±1.5dB
Mean (Relative Gain)					
- at 5dBi	±0.3 (±0.8)	±3	±0.25 (±0.5)	±1 (±2)	Envelope
- at -5dBi	±0.9 (±1.85)	±3	±0.3 (±0.7)	±1.5 (±3)	only
- at -15dBi	several dB	±3	±0.5 (±1.7)	±2.5 (±5)	

Note: Estimates of worst case errors are given in Brackets.

5. SYSTEM CONSIDERATIONS

In the preceding chapter the emphasis was on the radiation pattern aspects of the antenna installation. Of equal importance to the operator are a variety of other aspects which are mentioned in the following. These perhaps less obviously important questions are related to the system and its interaction with the whole airframe.

5.1 AERODYNAMICS

The traditional form of low gain aircraft antenna is the blade. This structure lends itself happily to the construction of antennas working in the frequency range 30 MHz to at least 10 GHz. In the lower reaches of this range they are physically quite large and can have impact on the stability of the airframe. This can be of use on occasion but is more likely to be a handicap on highly manoeuvrable aircraft.

An important consequence of the aerodynamic effects is that quite large side loads will be generated during manoeuvre. This can require airframe modifications prior to flight clearance of the installation.

5.2 CABLING

The system requirement in terms of the signal attenuation tolerable between the antenna and avionics is a continuing problem in many aircraft. The achievable pattern performance has been known to be dictated by this factor as a consequence of a poor site being the only option within this limit.

A common trap for the unwary in this respect is the number of cable breaks which may be necessary for cable installation and which can add substantially to the overall loss.

5.3 COUPLING

The antenna will by its nature interact with all other antennas on the airframe. The siting study should ensure that sufficient physical spacing can be provided to minimise this but for some systems this will be impossible.

Coupling problems fall into two categories:

- Radiation pattern distortion of the primary antenna by a parasitic element from another system.

This is principally physical obscuration and can, for blade antennas, be minimised by a separation of 1 metre or 4 wavelengths, whichever is the lesser. Coupling is of particular concern in homing systems which rely on controlled pattern shapes to deduce the desired information, ILS is a case in point.

- Coupling into the parasitic element causing performance reductions or even failure of the victim system.

High levels of pickup can cause saturation of the victim equipment when the system under consideration radiates leading to reduced sensitivity or blanking of the received signal. This is particularly a problem for high pulse power systems such as IFF (in civil terms ATC/SSR). In some instances such systems may have to produce pulses to trigger protection devices in the victim equipments to prevent erroneous data being gathered.

The first problem can only be resolved by radiation pattern measurement, as described earlier, at an early stage using scale models. The second can also be addressed by model measurement or by mathematical techniques these are normally based on empirically derived relationships.

5.4 RADAR CROSS SECTION

The echoing area of an airframe is often of great interest to the system designer and significant effort is often directed to its control. An antenna is designed to interact with incident electro-magnetic fields and as a consequence will always contribute significantly to the overall radar cross section. This effect can be quite startling for antennas illuminated by frequencies above their nominal operating frequency.

5.5 ELECTRO-MAGNETIC COMPATIBILITY

The aim of an antenna system is to radiate efficiently and when transmitting this can result in very large currents flowing on the airframe.

The sensitive electronics incorporated in modern avionics has proven on occasion sensitive to these currents and the airframe system authority must bear these effects in mind when specifying transmitter powers and the routing of cables from antennas or between avionic equipments.

The implications of these considerations will become broader as the trend to reinforced plastics structures and "fly by wire" systems continues.

6. CONCLUSIONS

The treatment given to many aspects touched upon in this discussion can only be brief in view of the breadth of the items to be considered. The attempt has been made to show the range of consideration needed in the practical design of an antenna installation and where theoretical techniques can be applied.

The aim of the antenna installer is always wider than the simple placing of a radiator on an aircraft and it is in the peripheral areas that so many interesting effects can be seen.

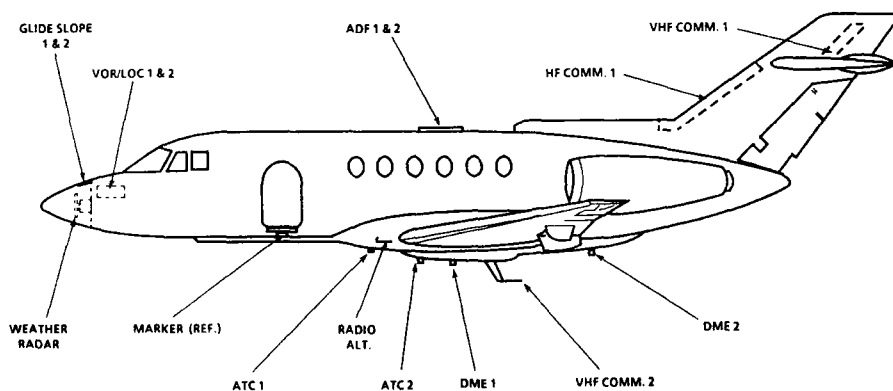


Figure 1 LOCATION OF ANTENNAS ON BAe 125

BIBLIOGRAPHY

This bibliography with abstracts has been prepared to support AGARD Lecture Series No.165 by Director Scientific Information Services (DSIS), Canada in consultation with the Lecture Series Director, Dr. J.S. Belrose, Communications Research Centre, Ottawa. The references contained herein are from the DSIS and NTIS databases, restricted to the years 1983 to 1989. They are in order by date, most recent first. Document requests should be directed to the defence information services of the requestor's own country. Please note that certain documents may not be released to all requestors.

DSIS DATABASE

88-00102 UNCLASSIFIED UNLIMITED

(U) A SYSTEM PERFORMANCE FEASIBILITY STUDY FOR A SANDWICH WIRE FREQUENCY SQUINTING RADAR ARRAY ANTENNA

Communications Research Centre, Ottawa ONT (CAN)
CRC-1411 Jun 87 50 pages
Chan, H.C.

(U) Sandwich Wire Antenna Arrays are attractive for its light-weight and low-cost attributes. However, this class of antennas have not been used in radar systems due to its moderate power handling capability and a lack of satisfactory design procedures for low-sidelobe applications. Recently, there has been interest in employing the Sandwich Wire Antenna in surveillance radar system on small naval vessels. Design studies for sandwich wire antennas have been undertaken. Being a travelling wave antenna, the main beam of this antenna possesses a frequency dependent squint. There is concern that this squint may have adverse effect on the signal processing system designed for radar systems employing conventional antennas. In this report, a feasibility study is carried out to evaluate the effects of the frequency squint on existing radar signal processing performance. In addition, new modes of operation designed to exploit the frequency squint are also identified and evaluated.

87-03960 UNCLASSIFIED UNLIMITED

(U) ADAPTIVE ANTENNAS

STC Defence Systems, Paignton (England)
AGARD-LS-151-PAP-6 Apr 87 21 pages
Barton, P.

@Lecture Series sponsored by Avionics Panel & Consultant & Exchange Programme, May 1987, in Italy, Germany, Turkey; CONTAINED IN 87-03966

(U) The basic principles of adaptive antennas are outlined in terms of the Wiener-Hopf expression for maximising signal to noise ratio in an arbitrary noise environment; the analogy with generalised matched filter theory provides a useful aid to understanding. For many applications, there is insufficient information to achieve the above solution and thus non-optimum constrained null steering algorithms are also

described, together with a summary of methods for preventing wanted signals being nulled by the adaptive system. The three generic approaches to adaptive weight control are then discussed; correlation steepest descent, weight perturbation and direct solutions based on sample matrix conversion. The tradeoffs between hardware complexity and performance in terms of null depth and convergence rate are outlined. The sidelobe cancellor technique is described, concentrating on large aperture main antennas associated with a small number of auxiliary antennas.

87-03963 UNCLASSIFIED UNLIMITED

(U) AIRCRAFT ANTENNAS/CONFORMAL ANTENNAS, MISSILE ANTENNAS

AEG Telefunken, Ulm (West Germany) Radio and Radar Systems Div
AGARD-LS-151-PAP-3 Apr 87 16 pages
Solbach, K.

@Lecture Series sponsored by Avionics Panel & Consultant & Exchange Programme, May 1987, in Italy, Germany, Turkey, CONTAINED IN 87-03966

(U) The lecture covers three major areas of airborne microwave antennas, namely: (1) Missile conformal telemetry/telecommand and radar-fuze antennas; (2) Missile and aircraft nose radar flatplate antennas; and (3) Aircraft Electronic Warfare antennas.

87-03964 UNCLASSIFIED UNLIMITED

(U) AIRCRAFT RADAR ANTENNAS

Westinghouse Electric Corp, Baltimore MD (US)
AGARD-LS-151-PAP-2 Apr 87 7 pages
Schrack, H.E.

@Lecture Series sponsored by Avionics Panel & Consultant & Exchange Programme, May 1987, in Italy, Germany, Turkey, CONTAINED IN 87-03966

(U) Many changes have taken place in airborne radar antennas since their beginnings over forty years ago. This lecture includes a brief historical overview of the advances in technology from mechanically scanned reflectors to modern multiple-function phased arrays. However, emphasis is not on history but on the "state-of-the-art" technology and trends for future airborne radar

systems. The status of rotating surveillance antennas is illustrated by the AN/APY-2 Airborne Warning and Control System (AWACS) slotted waveguide array, which achieved a significant breakthrough in sidelobe suppression. Gimballed flat-plate arrays in nose radomes are typified by the AN/APG-66 (F-16) antenna. Multifunction phased arrays are represented by the Electronically Agile Radar (EAR) antenna, which has achieved significant advances in performance versatility and reliability. Trends toward active aperture, adaptive, and digital beamforming arrays are briefly discussed.

87-03965 UNCLASSIFIED UNLIMITED

(U) BASIC PARAMETERS OF ANTENNAS FOR AIRCRAFT, SATELLITES AND MISSILES

Rome Air Development Center, Hanscom AFB MA (US)

AGARD-LS-151-PAP-1 Apr 87 17 pages
Mailloux, R.J.

@Lecture Series sponsored by Avionics Panel & Consultant & Exchange Programme, May 1987, in Italy, Germany, Turkey; CONTAINED IN 87-03966

(U) System requirements for airborne, satellite and missile antennas continue to place increasingly severe demands upon antenna technology. In general these requirements push toward the increased capability to control and modify antenna patterns, and away from the use of small antennas with broad radiation patterns. Increased control can imply several levels of added sophistication. At the lowest level it implies mechanical or electronic scanning of an antenna directive pattern, at the next level there are needs to produce precise low sidelobe radiation patterns, and at the highest level of complexity there is the need to actively suppress jammer interference through the use of adaptive control of a full array or an antenna with sidelobe cancellers. In addition to increased control, there is also a trend toward higher frequencies, even to EHF frequencies where arrays of several thousand elements are necessary for some applications.

87-03966 UNCLASSIFIED UNLIMITED

(U) MICROWAVE ANTENNAS FOR AVIONICS

Advisory Group for Aerospace Research and Development, Neuilly-sur-Seine (France) Avionics Panel

AGARD-LS-151 Apr 87 161 pages

@Lecture Series sponsored by Avionics Panel and the Consultant & Exchange Programme, May 1987, in Italy, Germany, Turkey

(No abstract)

87-03961 UNCLASSIFIED UNLIMITED

(U) MILLIMETER WAVE ANTENNAS FOR AVIONICS

Army Communications-Electronics Command, Fort Monmouth NJ (US)

AGARD-LS-151-PAP-5 Apr 87 8 pages

Schwering, F.K.

@Lecture Series sponsored by Avionics Panel & Consultant & Exchange Programme, May 1987, in Italy, Germany, Turkey; CONTAINED IN 87-03966

(U) An overview over the area of mm-wave antennas is presented with emphasis on possible avionics applications. For the purpose of the review, mm-wave radiating structures are grouped into two classes, i.e., antennas of conventional configuration and antennas based on new design concepts. The first class is composed of well known antennas such as reflector, lens, horn and slotted wave-guide antennas. The design principles and performance characteristics of these antennas are well established at microwave and lower frequencies and scaling into the mm-wave region is straightforward in most cases. The reduction in size and the tighter fabrication tolerances associated with the mm-wave region suggest certain modifications in the design and application of these antennas. But, in principle, they can be regarded as well understood. Most mm-wave antennas which are currently in use belong to this class of "conventional" antennas.

87-03958 UNCLASSIFIED UNLIMITED

(U) PRINTED CIRCUIT ANTENNA TECHNOLOGY

Massachusetts Univ, Amherst MA (US) Dept of Electrical and Computer Engineering

AGARD-LS-151-PAP-8 Apr 87 15 pages
Schaubert, D.H.

@Lecture Series sponsored by Avionics Panel & Consultant & Exchange Programme, May 1987, in Italy, Germany, Turkey; CONTAINED IN 87-03966

(U) Printed circuit antennas are being used on a variety of spacecraft, aircraft, and projectiles where their low profile and conformability minimize interference with the structural and aerodynamic properties of the vehicle. The basic microstrip antenna has been studied extensively, but some important aspects of its performance are not yet characterized by the analyses that are available to the designer. Furthermore, new applications in millimeter wave monolithic phased arrays are forcing the designer to use substrates that are electrically thick and that may have a high permittivity. The paper describes several characteristics of microstrip antennas and identifies some that are being studied with the aim of improving antenna performance. A brief description of analysis techniques available to the designer and researcher will be presented. Finally, several aspects of monolithic and

integrated phased arrays will be considered.

89-01014 UNCLASSIFIED DEFENCE PURPOSE ONLY
(U) NEAR VERTICAL INCIDENCE SKYWAVE FOR TACTICAL COMMUNICATIONS

Naval Ocean Systems Center, San Diego CA (US)
NOSC-TD-913 Jan 87 67 pages
AD-8111-792L

Dilley, D.M.

(U) Determine what procedural changes are required to employ near vertical incidence skywave (NVIS) for tactical land communications. Demonstrate effectiveness of NVIS in simulated field situations. Determine what equipment changes are required to employ NVIS techniques in the field and high frequency direction finding (HFDF) susceptibility of NVIS transmissions.

87-03350 UNCLASSIFIED DEFENCE PURPOSE ONLY
(U) THE PC VERSION OF THE MODEL CODE

Harry Diamond Labs, Adelphi MD (US)
HDL-TM-86-1 Feb 86 23 pages
AD-8100-327

Wilson, C.H.

(U) The report documents the IBM PC version of the MODEL CODE computer code. The modifications that transport MODEL CODE from the IBM mainframe to the IBM Personal Computer (PC-XT) are detailed. The MODEL CODE is used to obtain a quick estimate of the nuclear surface burst source-region electromagnetic pulse (EMP) environment and of the coupling effects on simple antennas and transmission lines. MODEL CODE will calculate and then plot the current and voltage waveforms coupled to a monopole (base-loaded) antenna or an end-fire-oriented transmission line. MODEL CODE will also plot the environment waveforms for the horizontal (azimuthal) magnetic fields, vertical and horizontal (radial) electric fields, air conductivity, radial Compton current, free electron density, and rate of ionization of air molecules as a function of time at specific fixed ranges, or peak values of the environments calculated as a function of range, bomb yield, atmospheric water vapor content, or air density.

85-04077 UNCLASSIFIED UNLIMITED
(U) DEVELOPMENT OF A PROTOTYPE REMOTELY PILOTED VEHICLE ANTENNA/ANTENNA MATCHING UNIT

Til-Tek Ltd, Kemptonville ONT (CAN)
Defence Research Establishment Ottawa, Ottawa
May 85 99 pages

Tilston, S.

Tilston, W.V.

Towne, J.

Tralman, T.

(U) The report covers work performed by TIL-TEK Limited from January 1985 through to June 1985 on

the development of a remotely piloted vehicle antenna/antenna matching unit. During this period, various manufacturers were contacted in the hope of finding a commercially available antenna that reasonably satisfied both the electrical and mechanical specifications.

85-02719 UNCLASSIFIED DEFENCE PURPOSE ONLY
(U) ANTENNAS FOR MOBILE RADIO SETS

Army Foreign Science and Technology Center, Charlottesville VA (US)

FSTC-HT-1456-84 Jan 85 4 pages
AD-8090 088

Troshin, I.

(U) The translation of a report written in Russian discusses various type of antennas for short-wave and ultrashort-wave portable radio sets and stations which are mounted on a transport base of various types. Portable stations in motion operate, as a rule, on collapsible-whip antennas with a length of not more than 1.5-1.8 m and having different designs (one-piece, telescopic, sectional, or collapsible). Even on the stand, more efficient wire antennas of the "sloped beam" and "symmetrical dipole" types. Portable ultrashort-wave radio sets operate on the stand on disc and biconical dipoles (for nondirectional communications) or on vertical semirhombic and collapsible log-periodic antennas. Mobile short-wave and ultrashort-wave radio sets also operate in motion on collapsible-whip antennas.

83-03592 UNCLASSIFIED UNLIMITED

(U) THE PERFORMANCE OF ANTENNAS IN THEIR OPERATIONAL ENVIRONMENT

Advisory Group for Aerospace Research and Development, Neuilly-sur-Seine (France)

AGARD-LS-131 Sep 83 141 pages

(U) Antennas can provide gain in any direction, and the effect of the operating environment on this parameter is fundamentally important for the performance of radio systems. Yet the effect of the environment of antennas is often overlooked. The performance of the antennas is more usually specified in terms of its operation over a perfectly conducting flat ground plane. The lecture series covers: techniques for measurement/prediction (numerical and experimental modelling); performance of fixed and transportable antennas (terrain effects, masts and buildings effects, re-radiation by supporting towers); performance of mobile antennas (effects of supporting platforms such as aircraft, ships and automobiles); performance of antennas in plasmas.

NTIS DATABASE

1376954 N89-15324/1/XAB

NASA-JSC (National Aeronautics and Space Administration-Johnson Space**Center) Antenna Near-Field Measurement System**

Cooke, W. P. ; Friederich, P. G. ; Jenkins, B. M. ; Jameson, C. R. ; Estrada, J. P.

Georgia Inst. of Tech., Atlanta.

Report No.: NAS 1.26:172106; NASA-CR-172106

Oct 88 277p

Languages: English

Journal Announcement: GRAI8910; STAR2707

NTIS Prices: PC A13/MF A01

Country of Publication: United States

Work was completed on the near-field range control software. The capabilities of the data processing software were expanded with the addition of probe compensation. In addition, the user can process the measured data from the same computer terminal used for range control. The design of the laser metrology system was completed. It provides precise measurement of probe location during near-field measurements as well as position data for control of the translation beam and probe cart. A near-field range measurement system was designed, fabricated, and tested.

1345192 AD-A197 364/3/XAB

Design of Antenna Matching Networks Using a Microcomputer

Li, S. T. ; Tam, D. W.

Naval Ocean Systems Center, San Diego, CA.

May 88 2p

Languages: English

Journal Announcement: GRAI8824

NTIS Prices: PC A02/MF A01

Country of Publication: United States

In a communication system one of the primary concerns is maximum efficiency in signal transmission and reception. For maximum power transfer there must be an impedance match between the antenna and transmitter (or receiver). A basic problem is to design a coupling network between a given source and a given load so that the transfer of power from the source to the load is maximized over a given frequency band of interest. The device used to perform this impedance matching is called an antenna matching network. The number of pieces of equipment requiring an antenna connection in some applications may exceed the number of acceptable locations available for antennas. One solution is the use of broadband antennas that have a low VSMR (voltage standing-wave ratio) over the operating band. Broadband antennas are used in conjunction with multicouplers (filters

with multiple inputs) to provide a sufficient number of antenna connections. Despite the aid of Smith Charts, the traditional design of an antenna matching network by engineering experience and manual calculation means is an extremely time-consuming task. The paper is intended to relieve the engineer of the tedious numerical calculation involved in the network design. (KR)

1301322 AD-A187 600/2/XAB

Design of Impedance-Matching Networks for Broadband Antennas

Li, S. T. ; Tam, D. W.

Naval Ocean Systems Center, San Diego, CA.

Report No.: NOSC/TD-1148

Sep 87 84p

Languages: English

Journal Announcement: GRAI8808

NTIS Prices: PC A05/MF A01

Country of Publication: United States

An interactive BASIC language computer program to aid in the design of a matching network for a broadband antenna is described. Guidelines for the design of a broadband matching network are presented. The Antenna Matching (ANTMAT) program provides a computer-aided design tool for determining network topology and component values of the networks. It improves the speed and accuracy of the broadband matching network design procedures. An optimization algorithm finds that the values of the components that minimize the input reflection coefficient. At first, the optimization algorithm with an exponential weighting function is employed for determining a list of network candidates (either a pi network for a T network) from which a network topology is selected. After a topology is specified, the optimization algorithm with other weighting function is used for finding optimum element values.

1299640 N88-12498/7/XAB

New Main Reflector, Subreflector and Dual Chamber Concepts for Compact Range Applications

Pistorius, C. W. I. ; Burnside, W. D.

Ohio State Univ., Columbus.

Report No.: NAS 1.26:181506; OSU-716148-22; NASA-CR-181506

Aug 87 350p

Languages: English

Journal Announcement: GRAI8807; STAR2604

NTIS Prices: PC A15/MF A01

Country of Publication: United States

A compact range is a facility used for the measurement of antenna radiation and target scattering problems. Most presently

available parabolic reflectors do not produce ideal uniform plane waves in the target zone. Design improvements are suggested to reduce the amplitude taper, ripple and cross polarization errors. The ripple caused by diffractions from the reflector edges can be reduced by adding blended rolled edges and shaping the edge contour. Since the reflected edge continues smoothly from the parabola onto the rolled surface, rather than being abruptly terminated, the discontinuity in the reflected field is reduced which results in weaker diffracted fields. This is done by blending the rolled edges from the parabola into an ellipse. An algorithm which enables one to design optimum blended rolled edges was developed that is based on an analysis of the continuity of the surface radius of curvature and its derivatives across the junction. Furthermore, a concave edge contour results in a divergent diffracted ray pattern and hence less stray energy in the target zone. Design equations for three-dimensional reflectors are given. Various examples were analyzed using a new physical optics method which eliminates the effects of the false scattering centers on the incident shadow boundaries. A Gregorian subreflector system, in which both the subreflector and feed axes are tilted, results in a substantial reduction in the amplitude taper and cross polarization errors. A dual chamber configuration is proposed to eliminate the effects of diffraction from the subreflector and spillover from the feed. A computationally efficient technique, based on ray tracing and aperture integration, was developed to analyze the scattering from a lossy dielectric slab with a wedge termination.

1272135 DE87007353/XAB

Enhancements and Limitations of the Code NEC (Numerical Electromagnetics Code) for Modeling Electrically Small Antennas

Burke, G. J.
Lawrence Livermore National Lab., CA.
Report No.: UCID-20970
Jan 87 21p
Languages: English
Journal Announcement: GRAI8722; NSA1200
Portions of this document are illegible in microfiche products. Original copy available until stock is exhausted.

NTIS Prices: PC A02/MF A01

Country of Publication: United States

This report discusses modifications to improve precision in Numerical Electromagnetics Code (NEC) Method-of-Moments computer code for modeling antennas by the numerical solution of an integral equation. These modifications

appear to correct most of the problems with dipoles or other open wires at VLF. Remaining problems with small loops are discussed and the limitations in modeling small loops are demonstrated. A discussion of problems presented by the introduction of a ground plane into the model is also presented. (ERA citation 12:031418)

1272134 DE87007352/XAB

NEC (Numerical Electromagnetics Code)-HYBRID User's Guide

Burke, G. J.
Lawrence Livermore National Lab., CA.
Report No.: UCID-20959
Jan 87 11p
Languages: English
Journal Announcement: GRAI8722; NSA1200
Microfiche only, copy does not permit paper copy reproduction. Original copy available until stock is exhausted.

NTIS Prices: MF A01

Country of Publication: United States

The combination of Numerical Electromagnetics Code (NEC) Method-of-Moments and the geometrical theory of diffraction (GTD) provides a useful new modeling capability. A combination of these two techniques can provide an efficient method for modeling structures such as wire antennas near large conducting bodies. This approach has been used in NEC-HYBRID, and sample input data and results are described. (ERA citation 12:031417)

1274953 DE87007324/XAB

Finite Difference Approach to Microstrip Antenna Design

Barth, M. J. ; Bevensee, R. M. ; Pennock, S. T.
Lawrence Livermore National Lab., CA.
Report No.: UCRL-95872; CONF-870655-4
Dec 86 5p
Languages: English
Journal Announcement: GRAI8723; NSA1200
Antennas and Propagation Society symposium, Blacksburg, VA, USA, 15 Jun 1987.
Microfiche only, copy does not permit paper copy reproduction.

NTIS Prices: MF A01

Country of Publication: United States

Microstrip antennas have received increased attention in recent years, due to their size and cost advantages. Analysis of the microstrip structure has proved difficult due to the presence of the dielectric substrate, particularly for complex geometries. One possible approach to a solution is the use of a finite difference computer code to model a proposed microstrip antenna design. The models are easily constructed and altered, and code versions are

available which allow input impedance or far-field patterns to be calculated. Results for some simple antenna geometries will be presented. (ERA citation 12:033424)

1224370 N87-10225/7/XAB

Study of Microstrip Antenna Arrays and Related Problems

Lo, Y. T.

Illinois Univ. at Urbana-Champaign.

Report No.: NAS 1.26:179714; NASA-CR-179714

3 Oct 86 61p

Languages: English

Journal Announcement: GRAI8703; STAR2501

NTIS Prices: PC A04/MF A01

Country of Publication: United States

In February, an initial computer program to be used in analyzing the four-element array module was completed. This program performs the analysis of modules composed of four rectangular patches which are corporately fed by a microstrip line network terminated in four identical load impedances. Currently, a rigorous full-wave analysis of various types of microstrip line feed structures and patches is being performed. These tests include the microstrip line feed between layers of different electrical parameters. A method of moments was implemented for the case of a single dielectric layer and microstrip line fed rectangular patches in which the primary source is assumed to be a magnetic current ribbon across the line some distance from the patch. Measured values are compared with those computed by the program.

1214307 N86-28973/3/XABXABXABXAB

Near-Field Testing of the 15-Meter Model of the Hoop Column Antenna. Volume 3. Near- and Far-Field Plots for the JPL Feed

Hoover, J. ; Kefauver, M. ; Cencich, T. ; Osborn, J.

Martin Marietta Aerospace, Denver, CO. Denver Div.

Report No.: NAS 1.26:178061; MCR-85-640-V-3; NASA-CR-178061

Mar 86 147p

Languages: English

Journal Announcement: GRAI8624; STAR2420

NTIS Prices: PC A07/MF A01

Country of Publication: United States

Technical results from near-field testing of the 15-meter model of the hoop column antenna at the Martin Marietta Denver Aerospace facility are discussed. The antenna consists of a deployable central column and a 15 meter hoop, stiffened by cables into a structure with a high tolerance repeatable surface and offset

feed location. The surface has been configured to have four offset parabolic apertures, each about 6 meters in diameter, and is made of gold plated molybdenum wire mesh. Pattern measurements were made with feed system radiating at frequencies of 7.73, 11.60, 2.27, 2.225, and 4.26 (all in GHz). This report (Volume 3) gives the detailed patterns measured with the JPL feed (2.225 GHz). Volume 1 covers the testing from an overall viewpoint and contains information of generalized interest for testing large antennas, including the deployment of the antenna in the Martin Facility and the measurements to determine mechanical stability and trueness of the reflector surface, the test program outline, and a synopsis of antenna electromagnetic performance. A detailed listing of the antenna patterns for the LaRC feeds (7.3, 11.60, 2.27, and 4.26 GHz) are given in Volume 2 of this report.

1253520 N87-20763/5/XAB

Automatic Graphics Applied to Antenna Computer Aided Design

Biagi, M. ; Devincenzi, P.

Selenia S.p.A., Pomezia (Italy). Antennas Lab.

1986 19p

Languages: English

Journal Announcement: GRAI8715; STAR2513

In Its Rivista Tecnica Selenia, Volume 10, Number 1 p 32-49 (See N87-20759 13-61).

NTIS Prices: (Order as N87-20759 PC A04/MF A01)

Country of Publication: Italy

Antenna design programs in FORTRAN 4 which utilize the Tektronix Plot 10 system graphic routine are shown. The programs are available on Univac 1100/80 computer.

1288887 N88-10233/0/XAB

Formoptimierung von Linearantennen (Shape Optimization of Linear Antennas)

Gotthard, O.

Technische Univ. Muenchen (Germany, F.R.). Fakultät fuer Elektrotechnik.

Report No.: ETN-87-90413

1986 200p

Languages: German Document Type: Thesis

Journal Announcement: GRAI8803; STAR2601

Text in German.

NTIS Prices: PC A09/MF A01

Country of Publication: Germany, Federal Republic of

A method for the optimization of antenna geometry is presented. The arrangement of the antenna conductors in space is described by a sum of known shape functions. Target functions are formulated for the mathematical realization of the

required characteristics. The antenna shape optimization is reduced to the problem of the determination of those variables in the shape function with which the target function reaches its maximum value. The FORTRAN computer program AOPT for the optimization of linear antennas is enclosed.

1195340 DE86008108108146/XAB

Scale Model Evaluation of a Reflecting Antenna Design for an Indoor EMP Simulator

Hudson, H. G. ; Pennock, S. T.
Lawrence Livermore National Lab., CA.

Report No.: UCID-20685

Dec 85 30p

Languages: English

Journal Announcement: GRAI8616; NSA1100

Portions of this document are illegible in microfiche products. Original copy available until stock is exhausted.

NTIS Prices: PC A03/MF A01

Country of Publication: United States

The purpose of this effort was to construct a scale model of a cusp antenna and evaluate its performance. Measurements were taken to describe the distortion the presence of the cusp introduced into a free space incident field. Measurements of the hangar were also taken to serve as a reference point to gauge how much the cusp was able to reflect energy. As will be described in detail later, the measured performance of the scale model cusp antenna does not agree with the predicted computer solutions from EMA. 22 figs. (ERA citation 11:023310)

1276339 N87-27096/3/XAB

Airborne Antenna Radiation Pattern Code User's Manual

Burnside, W. D. ; Kim, J. J. ; Grandchamp, B. ;
Rojas, R. G. ; Law, P.

Ohio State Univ., Columbus.

Report No.: NAS 1.26:181249; REPT-716199-4;
NASA-CR-181249

Sep 85 261p

Languages: English

Journal Announcement: GRAI8723; STAR2521

NTIS Prices: PC A12/MF A01

Country of Publication: United States

The use of a newly developed computer code to analyze the radiation patterns of antennas mounted on a ellipsoid and in the presence of a set of finite flat plates is described. It is shown how the code allows the user to simulate a wide variety of complex electromagnetic radiation problems using the ellipsoid/plates model. The code has the capacity of calculating radiation

patterns around an arbitrary conical cut specified by the user. The organization of the code, definition of input and output data, and numerous practical examples are also presented. The analysis is based on the Uniform Geometrical Theory of Diffraction (UTD), and most of the computed patterns are compared with experimental results to show the accuracy of this solution.

1234036 N87-14566/0/XAB

Practical Guide to the Design of Rhombic Aerials

Tyler, J. N.

Royal Aircraft Establishment, Farnborough (England).

Report No.: RAE-TN-RAD-NAV-266; BR98240

Apr 85 39p

Languages: English

Journal Announcement: GRAI8708; STAR2506

NTIS Prices: PC A03/MF A01

Country of Publication: United Kingdom

Basic principles of the rhombic aerial are explained and design charts are presented. These, compiled using two simple computer programs, offer a considerable time saving in design. Examples of aerials designed using the charts are given, together with polar diagrams of their theoretical performance.

1280006 N87-27870/1/XAB

Simulation and Analysis of Airborne Antenna Radiation Patterns

Kim, J. J. ; Burnside, W. D.

Ohio State Univ., Columbus.

Report No.: NAS 1.26:181250; REPT-716199-1;
NASA-CR-181250

Dec 84 345p

Languages: English

Journal Announcement: GRAI8724; STAR2522

NTIS Prices: PC A15/MF A01

Country of Publication: United States

The objective is to develop an accurate and efficient analytic solution for predicting high frequency radiation patterns of fuselage-mounted airborne antennas. This is an analytic study of airborne antenna patterns using the Uniform Geometrical Theory of Diffraction (UTD). The aircraft is modeled in its most basic form so that the solution is applicable to general-type aircraft. The fuselage is modeled as a perfectly conducting composite ellipsoid; whereas, the wings, stabilizers, nose, fuel tanks, and engines, are simulated as perfectly conducting flat plates that can be attached to the fuselage and/or to each other. The composite-ellipsoid fuselage model is necessary to successfully simulate the wide variety of real

world fuselage shapes. Since the antenna is mounted on the fuselage, it has a dominant effect on the resulting radiation pattern so it must be simulated accurately, especially near the antenna. Various radiation patterns are calculated for commercial, private, and military aircraft, and the Space Shuttle Orbiter. The application of this solution to numerous practical airborne antenna problems illustrates its versatility and design capability. In most cases, the solution accuracy is verified by the comparisons between the calculated and measured data.

1222052 N86-32424/1/XAB

Theoretical and Computational Aspects of Adaptively Controlled Navstar Antennas

Pietersen, O. B. M. ; Klinker, F. ; Roefs, H. F. A. ; Lucas, A. M. G. M.
National Aerospace Lab., Amsterdam (Netherlands).

Report No.: NLR-TR-84004-U; B8665726

1 Jan 84 91p

Languages: English

Journal Announcement: GRAI8702; STAR2424

Sponsored by NATO.

NTIS Prices: PC A05/MF A01

Country of Publication: Netherlands

Jamming suppression antennas used by military users of Navstar Global Positioning System were studied. Literature on adaptive null-steering antennas was reviewed and a computer program was designed as a tool for the evaluation of antenna choices. The program is used to determine the requirements for threat suppression and compute the performance of null or beam-steering antennas in electronic warfare.

1089535 N84-32789/9

Characteristics and Capacities of the NASA Lewis Research Center High Precision 6.7- by 6.7-M Planar Near-Field Scanner

Sharp, G.R. ; Zakrajsek, R.J. ; Kunath, R.R. ; Raquet, C. A. ; Alexovich, R. E.

National Aeronautics and Space Administration, Cleveland, OH. Lewis Research Center.

Report No.: NAS 1.15:83785;E-2281; NASA-TN-83785 1984 27p

Languages: English

Journal Announcement: GRAI8426; STAR2222

Presented at the Meeting of the Antenna Meas. Tech. Assoc., San Diego, Calif., 2-4 Oct. 1984.

NTIS Prices: PC A03/MF A01

Country of Publication: United States

A very precise 6.7- by 6.7-m planar near-field scanner has recently become operational at

the NASA Lewis Research Center. The scanner acquires amplitude and phase data at discrete points over a vertical rectangular grid. During the design phase for this scanner, special emphasis was given to the dimensional stability of the structures and the ease of adjustment of the rails that determine the accuracy of the scan plane. A laser measurement system is used for rail alignment and probe positioning. This has resulted in very repeatable horizontal and vertical motion of the probe cart and hence precise positioning in the plane described by the probe tip. The resulting accuracy will support near-field measurements at 60 GHz without corrections. Subsystem design including laser, electronic and mechanical and their performance is described. Summary data are presented on the scan plane flatness and environmental temperature stability. Representative near-field data and calculated far-field test results are presented. Prospective scanner improvements to increase test capability are also discussed.

1051470 AD-P002 855/5

Computer Graphics Techniques for Aircraft EMC (Electromagnetic Compatibility) Analysis and Design

Kubina, S. J. ; Bhartia, P.

Defence Research Establishment, Ottawa Oct 83 9p

Languages: English

Journal Announcement: GRAI8412

This article is from 'Advanced Concepts for Avionics/Weapon System Design, Development and Integration: Conference Proceedings of the Avionics Panel Symposium (45th) Held at Ottawa, Canada on 18-22 April 1983,' AD-A138600, p22-1-22-9.

NTIS Prices: PC A02/MF A01

Country of Publication: Canada

This paper describes a comprehensive computer-aided system for the prediction of the potential interaction between avionics systems, with special emphasis on antenna-to-antenna coupling. The methodology is applicable throughout the life cycle of an avionic/weapon system, including system upgrades and retrofits. As soon as aircraft geometry and preliminary systems information becomes available, the computer codes can be used to selectively display proposed antenna locations, emitter/receptor response characteristics, electromagnetic interference (EMI) margins and the actual rayoptical paths of maximum antenna-antenna coupling for each potential interacting antenna set. The visibility of the entire interaction matrix produces an appreciation and awareness that had heretofore been

unavailable. Antennas can be interactively relocated by track-ball (or joystick) and the analysis repeated at will for optimization or installation design study purposes. The codes can significantly simplify the task of the designer/analyst in effectively identifying critical interactions among an overwhelming large set of potential ones. In addition, it is an excellent design, development and analysis tool which simultaneously identifies both numerically and pictorially the EMI interdependencies among subsystems.

1060737 N84-21794/2

Experimental Spherical Near Field Antenna Test Facility, Phase 3. Volume 2: Study of an Offset Reflector Antenna with a Sagging Feed Support

Frandsen, A.

TICRA A/S, Copenhagen (Denmark)

Sponsor: National Aeronautics and Space Administration, Washington, DC.

Report No.: S-179-01-V-2; ESA-CR(P)-1834-V-2

Jun 83 63p

Languages: English

Journal Announcement: GRAI8415; STAR2212

NTIS Prices: PC A04/MF A01

Country of Publication: Denmark

A numerical model for satellite antenna simulations based on calculating the near field from the antenna by geometrical diffraction theory and then transforming to the far field was developed. For an offset reflector with a circular aperture and a nonrigid feed support, it is demonstrated that the impact of the sagging feed depends on antenna mount and orientation with respect to the measurement grid. Predicted far field pattern cuts show pattern degradations for different degrees of feed rigidity. The influence of reflector edge taper is also shown. Simulations for an idealized model antenna show that stiffening of the feed support considerably improves pattern prediction.

1049990 N84-17460/6

Program for Computer-Controlled Impedance Measurement within the Microwave Range

Geekstorp, P.

Foersvarets Forskningsanstalt, Stockholm (Sweden)

Report No.: FOA-C-30330-E3

Jun 83 66p

Languages: Swedish

Journal Announcement: GRAI8411; STAR2208

In Swedish; English Summary.

NTIS Prices: PC A04/MF A01

Country of Publication: Sweden

A program for computer-controlled impedance measurement on antennas, by means of a network

analyzer, sweep generator, and frequency counter, is presented. The reflection factor was measured. The corresponding scattering parameter was determined in the form of its amount and phase from which the complex number S_{11} was obtained, corrected by parameters, and stored on a file specified by the user. The correction parameters were determined by a calibrating routine preceding the measurement. A minicomputer was used for control and monitoring of the measuring data. A program plotting the S-parameters in a Smith chart is shown.

0994390 AD-P001 101/5

Computer Aided Analysis of a Blass Feed Network for Wide Instantaneous Band Phased Array Application

Mojtowicz, J. ; Hacker, P. S. ; Ramsey, K. G.

Westinghouse Defense and Electronic Systems Center, Baltimore, MD. Systems Development Div.

Jan 83 24p

Languages: English

This article is from 'Proceedings of the Antenna Applications Symposium

(1982) Held at Illinois University at Urbana on 22-24 September 1982,'

AD-A129 356.

NTIS Prices: PC A02/MF A01

Country of Publication: United States

Engineering and next generation radar system studies have identified a need for phased array antennas with wide instantaneous bandwidths. Typical phased array antenna bandwidths which are compatible with these system requirements are in the range of 5 to 10%.

0993947 AD-A129 356/2

Proceedings of the Antenna Applications Symposium (1982) Held at Illinois

University at Urbana on 22-24 September 1982

Rome Air Development Center, Griffiss AFB, NY.

Languages: English Journal Announcement: GRAI8320

For sales information of individual items see AD-P001 095 - AD-P001 124.

NTIS Prices: PC A99/MF A01

Country of Publication: United States

The Proceedings of the 1982 Antenna Applications Symposium is a collection of the State-of-the-Art papers relating to Phased Array Antennas, Millimeter Wave Antennas, Microstrip and Conformal Antennas and Reflector Antennas. (Author)

REPORT DOCUMENTATION PAGE			
1. Recipient's Reference	2. Originator's Reference	3. Further Reference	4. Security Classification of Document
	AGARD-LS-165	ISBN 92-835-0526-3	UNCLASSIFIED
5. Originator	Advisory Group for Aerospace Research and Development North Atlantic Treaty Organization 7 rue Ancelle, 92200 Neuilly sur Seine, France		
6. Title	MODERN ANTENNA DESIGN USING COMPUTERS AND MEASUREMENT: APPLICATION TO ANTENNA PROBLEMS OF MILITARY INTEREST		
7. Presented on	19—20 October 1989 in Ankara, Turkey, on 23—24 October 1989 in Lisbon, Portugal and on 26—27 October 1989 in London, United Kingdom.		
8. Author(s)/Editor(s)		9. Date	
Various		September 1989	
10. Author's/Editor's Address		11. Pages	
Various		224	
12. Distribution Statement	This document is distributed in accordance with AGARD policies and regulations, which are outlined on the Outside Back Covers of all AGARD publications.		
13. Keywords/Descriptors			
<div style="display: flex; justify-content: space-between;"> <div> Antennas Design Computerized simulation </div> <div> Data processing Military communications </div> </div>			
14. Abstract			
<p>The working environment in which an antenna is installed may substantially modify such antenna parameters as radiation efficiency, impedance, bandwidth, power handling capacity, pattern, etc. The need for more accurate antenna design, combined with the continuing growth of computational techniques, are complementing the more traditional approaches of measurement and analysis to vastly broaden the breadth and depth of problems that are now quantifiable. Computers are being used not only for numerical modelling/simulation, but also for measurement, data acquisition, and subsequent transformation of data. The newly available computational techniques is changing the way we think about, formulate, solve and interpret problems.</p> <p>This Lecture Series, sponsored by the AGARD Electromagnetic Wave Propagation Panel, has been implemented by the Consultant and Exchange Programme.</p>			

<p>AGARD Lecture Series No.165 Advisory Group for Aerospace Research and Development, NATO MODERN ANTENNA DESIGN USING COMPUTERS AND MEASUREMENT: APPLICATION TO ANTENNA PROBLEMS OF MILITARY INTEREST Published September 1989 224 pages</p> <p>The working environment in which an antenna is installed may substantially modify such antenna parameters as radiation efficiency, impedance, bandwidth, power handling capacity, pattern, etc. The need for more accurate antenna design, combined with the continuing growth of computational techniques, is complementing the more traditional approaches of measurement and analysis to</p> <p>P.T.O.</p>	<p>AGARD-LS-165</p> <p>Antennas Design Computerized simulation Data processing Military communication</p>	<p>AGARD Lecture Series No.165 Advisory Group for Aerospace Research and Development, NATO MODERN ANTENNA DESIGN USING COMPUTERS AND MEASUREMENT: APPLICATION TO ANTENNA PROBLEMS OF MILITARY INTEREST Published September 1989 224 pages</p> <p>The working environment in which an antenna is installed may substantially modify such antenna parameters as radiation efficiency, impedance, bandwidth, power handling capacity, pattern, etc. The need for more accurate antenna design, combined with the continuing growth of computational techniques, is complementing the more traditional approaches of measurement and analysis to</p> <p>P.T.O.</p>	<p>AGARD-LS-165</p> <p>Antennas Design Computerized simulation Data processing Military communication</p>
<p>AGARD Lecture Series No.165 Advisory Group for Aerospace Research and Development, NATO MODERN ANTENNA DESIGN USING COMPUTERS AND MEASUREMENT: APPLICATION TO ANTENNA PROBLEMS OF MILITARY INTEREST Published September 1989 224 pages</p> <p>The working environment in which an antenna is installed may substantially modify such antenna parameters as radiation efficiency, impedance, bandwidth, power handling capacity, pattern, etc. The need for more accurate antenna design, combined with the continuing growth of computational techniques, is complementing the more traditional approaches of measurement and analysis to</p> <p>P.T.O.</p>	<p>AGARD-LS-165</p> <p>Antennas Design Computerized simulation Data processing Military communication</p>	<p>AGARD Lecture Series No.165 Advisory Group for Aerospace Research and Development, NATO MODERN ANTENNA DESIGN USING COMPUTERS AND MEASUREMENT: APPLICATION TO ANTENNA PROBLEMS OF MILITARY INTEREST Published September 1989 224 pages</p> <p>The working environment in which an antenna is installed may substantially modify such antenna parameters as radiation efficiency, impedance, bandwidth, power handling capacity, pattern, etc. The need for more accurate antenna design, combined with the continuing growth of computational techniques, is complementing the more traditional approaches of measurement and analysis to</p> <p>P.T.O.</p>	<p>AGARD-LS-165</p> <p>Antennas Design Computerized simulation Data processing Military communication</p>

<p>vastly broaden the breadth and depth of problems that are now quantifiable. Computers are being used not only for numerical modelling/simulation, but also for measurement, data acquisition, and subsequent transformation of data. The newly available computational techniques are changing the way we think about, formulate, solve and interpret problems.</p> <p>This Lecture Series, sponsored by the AGARD Electromagnetic Wave Propagation Panel, has been implemented by the Consultant and Exchange Programme.</p> <p>ISBN 92-835-0526-3</p>	<p>vastly broaden the breadth and depth of problems that are now quantifiable. Computers are being used not only for numerical modelling/simulation, but also for measurement, data acquisition, and subsequent transformation of data. The newly available computational techniques are changing the way we think about, formulate, solve and interpret problems.</p> <p>This Lecture Series, sponsored by the AGARD Electromagnetic Wave Propagation Panel, has been implemented by the Consultant and Exchange Programme.</p> <p>ISBN 92-835-0526-3</p>
<p>vastly broaden the breadth and depth of problems that are now quantifiable. Computers are being used not only for numerical modelling/simulation, but also for measurement, data acquisition, and subsequent transformation of data. The newly available computational techniques are changing the way we think about, formulate, solve and interpret problems.</p> <p>This Lecture Series, sponsored by the AGARD Electromagnetic Wave Propagation Panel, has been implemented by the Consultant and Exchange Programme.</p> <p>ISBN 92-835-0526-3</p>	<p>vastly broaden the breadth and depth of problems that are now quantifiable. Computers are being used not only for numerical modelling/simulation, but also for measurement, data acquisition, and subsequent transformation of data. The newly available computational techniques are changing the way we think about, formulate, solve and interpret problems.</p> <p>This Lecture Series, sponsored by the AGARD Electromagnetic Wave Propagation Panel, has been implemented by the Consultant and Exchange Programme.</p> <p>ISBN 92-835-0526-3</p>

AGARD

NATO OTAN

7 rue Ancelle • 92200 NEUILLY-SUR-SEINE
FRANCE

Telephone (1)47.38.57.00 • Telex 610 176

**DISTRIBUTION OF UNCLASSIFIED
AGARD PUBLICATIONS**

AGARD does NOT hold stocks of AGARD publications at the above address for general distribution. Initial distribution of AGARD publications is made to AGARD Member Nations through the following National Distribution Centres. Further copies are sometimes available from these Centres, but if not may be purchased in Microfiche or Photocopy form from the Purchase Agencies listed below.

NATIONAL DISTRIBUTION CENTRES

BELGIUM

Coordonnateur AGARD — VSL
Etat-Major de la Force Aérienne
Quartier Reine Elisabeth
Rue d'Evere, 1140 Bruxelles

LUXEMBOURG

See Belgium

NETHERLANDS

Netherlands Delegation to AGARD
National Aerospace Laboratory, NLR
P.O. Box 126
2600 AC Delft

CANADA

Director Scientific Information Services
Dept of National Defence
Ottawa, Ontario K1A 0K2

NORWAY

DENMARK

Danish Defence
Ved Idraetst
2100 Copenhagen



National Aeronautics and
Space Administration

Washington, D.C. 20546 **SPECIAL FOURTH CLASS MAIL
BOOK**

Postage and Fees Paid
National Aeronautics and
Space Administration
NASA-451

Official Business
Penalty for Private Use \$300



FRANCE

O.N.E.R.A.
29 Avenue de
92320 Châtillon

GERMANY

Fachinforma-
Physik, Math
Karlsruhe
D-7514 Egg

L1 001 AGARDL3165891127S002672D
DEPT OF DEFENSE
DEFENSE TECHNICAL INFORMATION CENTER
DTIC-FDAC
CAMERON STATION BLDG 5
ALEXANDRIA VA 223046145

GREECE

Hellenic Air
Aircraft Sup
Department
Holargos, Athens

ICELAND

Director of Aviation
c/o Flugrad
Reykjavik

Kentigern House
65 Brown Street
Glasgow G2 8EX

ITALY

Aeronautica Militare
Ufficio del Delegato Nazionale all'AGARD
3 Piazzale Adenauer
00144 Roma/EUR

UNITED STATES

National Aeronautics and Space Administration (NASA)
Langley Research Center
M/S 180
Hampton, Virginia 23665

THE UNITED STATES NATIONAL DISTRIBUTION CENTRE (NASA) DOES NOT HOLD
STOCKS OF AGARD PUBLICATIONS, AND APPLICATIONS FOR COPIES SHOULD BE MADE
DIRECT TO THE NATIONAL TECHNICAL INFORMATION SERVICE (NTIS) AT THE ADDRESS BELOW.

PURCHASE AGENCIES

National Technical
Information Service (NTIS)
5285 Port Royal Road
Springfield
Virginia 22161, USA

ESA/Information Retrieval Service
European Space Agency
10, rue Mario Nikis
75015 Paris, France

The British Library
Document Supply Centre
Boston Spa, Wetherby
West Yorkshire LS23 7BQ
England

Requests for microfiche or photocopies of AGARD documents should include the AGARD serial number, title, author or editor, and publication date. Requests to NTIS should include the NASA accession report number. Full bibliographical references and abstracts of AGARD publications are given in the following journals:

Scientific and Technical Aerospace Reports (STAR)
published by NASA Scientific and Technical
Information Branch
NASA Headquarters (NIT-40)
Washington D.C. 20546, USA

Government Reports Announcements (GRA)
published by the National Technical
Information Services, Springfield
Virginia 22161, USA



Printed by Specialised Printing Services Limited
40 Chigwell Lane, Loughton, Essex IG10 3TZ

ISBN 92-835-0526-3